

XX (LEXI-) LEXICON GENETICS INC.
 PA XX
 PI Yu X, Miranda M, Turner CA;
 XX DR WPI; 2002-559797/64.
 DR N-PSDB; ABQ77623.

XX Polynucleotides encoding human lipases that are structurally related to animal lipases, particularly phospholipase B, useful for drug screening, diagnosis and in gene therapy of biological disorders -
 PT Claim 2; Page 37-41; 44pp; English.

XX The invention relates to a novel human lipase (NHL; ABB09555, ABB09556) and to nucleic acids encoding it (ABQ77623, ABQ77624). The NHL has structural similarity with animal lipases, particularly phospholipase B. Polynucleotides encoding NHL were obtained using human genomic sequences in conjunction with human thyroid and brain cDNA. The NHL gene is located on chromosome 2, and contains a C/T polymorphism at position 3953 of the open reading frame (ORF) resulting in an Ala/Val substitution at position 1318 in the protein. NHL nucleotides and proteins are useful for treating disorders such as inflammatory or proliferative disease, infectious disease, clotting disorders and cancer. They can also be used in screening for compounds useful in the treatment of mental, biological or medical disorders, as diagnostic reagents, in clinical trial monitoring and in cosmetic and nutraceutical applications. NHL nucleotides can additionally be used in the detection of disease-associated mutations, in the analysis of gene expression, for the recombinant expression of NHL, to generate transgenic animals, in gene therapy, and as part of ribozyme and/or triple helix sequences useful in the modulation of NHL gene expression. The present sequence represents the Ala 1318 variant of NHL.

XX Sequence 1458 AA;

Query Match 100.0% Score 776; DB 23; Length 1458;
 Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 1458; Conservative 0;

Qy . 1 MGJLRPGIFILELLILLGQCTPQHTSPRKSRLTGCOLWBTLNQSPFPCPNPNUKGVNNPSK 60
 Db 1 MGJLRPGIFILELLILLGQCTPQHTSPRKSRLTGCOLWBTLNQSPFPCPNPNUKGVNNPSK 60
 Qy . 61 SVMSIKPEDIKVAAIGNEIIPDGGTCDLEKDWTPEQQVONGWMTLSPIRYSPS 120
 Db 61 SVHSIKPDSIKFVAIGNEIIPDGGTCDLEKDWTPEQQVONGWMTLSPIRYSPS 120
 Qy . 121 VPMVPUCHTKRVTPHDGAEDMTOAQEJVRNMKENLQDFQDWKLINFFNSAQCYL 180
 Db 121 VPMPVUCHTKRVTPHDGAEDMTOAQEJVRNMKENLQDFQDWKLINFFNSAQCYL 180
 Qy . 181 PSAQONGLAAGGYDDELMGVDYILOQEVRAFPVNLVDISEVAEVSRQMQHTWSPAPERCN 240
 Db 181 PSAQONGLAAGGYDDELMGVDYILOQEVRAFPVNLVDISEVAEVSRQMQHTWSPAPERCN 240
 Qy . 241 CSBETTRILAKVNMWSOBANWLLASSRYSBESFTVQPFYETTPSLSEDPRLD 300
 Db 241 CSBETTRILAKVNMWSOBANWLLASSRYSBESFTVQPFYETTPSLSEDPRLD 300
 Qy . 301 STTLAWHWNRMNMPAGEKDEPLSVKGRPMKPSQSPSYLPSYRNSYLTQKOPDKL 360
 Db 301 STTLAWHWNRMNMPAGEKDEPLSVKGRPMKPSQSPSYLPSYRNSYLTQKOPDKL 360
 Qy . 361 EYREGAEATRCPPDDPSDTPVPTSVARLKCADINTGALGDSITAGNGAGSTPENVLVITQ 420
 Db 361 EYREGAEATRCPPDDPSDTPVPTSVARLKCADINTGALGDSITAGNGAGSTPENVLVITQ 420
 Qy . 421 YRGLSWSVGGDENITGTVTNLIREFPLSKGSPVSGKGKESPNAFLNOAVAGGRAEL 480
 Db 421 YRGLSWSVGGDENITGTVTNLIREFPLSKGSPVSGKGKESPNAFLNOAVAGGRAEL 480
 481 PVOARRLVDLMKNDTRHFQDMKILTEPIGNGLCDFCNDLVHYSQNFTNIGKALDI 540

Do 481 PVGARRLVDLMKNDTRHFQDMKILTEPIGNGLCDFCNDLVHYSQNFTNIGKALDI 540
 Qy 541 LHAEVPRAFVNLYVTEI VNRLBYQKVKYCPMRILSLCPCVULKEDDNSTBLATIEP 600.
 Do 541 LHAEVPRAFVNLYVTEI VNRLBYQKVKYCPMRILSLCPCVULKEDDNSTBLATIEP 600
 Do 601 NKCFQEKTHOLIESGRYVTRDFTVWQPFENNMDPRTSEGLPDNSPFAPCFHSSKS 660
 Do 601 NKCFQEKTHOLIESGRYVTRDFTVWQPFENNMDPRTSEGLPDNSPFAPCFHSSKS 660
 Qy 651 HSRASALWNMMLEPVGQKTTHKFENKINTTCPNOVOPFLRYKNSMQGHGTWLPCCR 720
 Do 651 HSRASALWNMMLEPVGQKTTHKFENKINTTCPNOVOPFLRYKNSMQGHGTWLPCCR 720
 Qy 721 APGLAHPTSVHAIKPADITQVWAALGDSITAGNGIGSKPDDLPVTTQYRGLSYSSAGGS 780
 Do 721 APGLAHPTSVHAIKPADITQVWAALGDSITAGNGIGSKPDDLPVTTQYRGLSYSSAGGS 780
 Qy 781 LENVTTLNILRBRPNRNTGAVGVTGDRDNTNAPLNQRPVGAKAEDLMSQVOTLQKND 840
 Do 781 LENVTTLNILRBRPNRNTGAVGVTGDRDNTNAPLNQRPVGAKAEDLMSQVOTLQKND 840
 Qy 841 DHRNFHDWKVITVLLGSDLCYDCHTSNLYSAANFDHNRALDVREPVRLV 900
 Do 841 DHRNFHDWKVITVLLGSDLCYDCHTSNLYSAANFDHNRALDVREPVRLV 900
 Qy 901 DFIAAPTMRQVEGANPDCKPVCQVQASVLCNCVTLRENSQELARBAFSAVRSRSMRLEV 960
 Do 901 DFIAAPTMRQVEGANPDCKPVCQVQASVLCNCVTLRENSQELARBAFSAVRSRSMRLEV 960
 Qy 961 SGYDQTDFSVLQLPFFQFQONIQLPVLAQGLPDTFFARDCHPNQKFISQALARALWML 1020
 Do 961 SGYDQTDFSVLQLPFFQFQONIQLPVLAQGLPDTFFARDCHPNQKFISQALARALWML 1020
 Do 1021 EPGSKTEFLDRAEMPTCPIONEPTRPENSYTPIKA1ENNGSDEFLCTEWKSN 1080
 Do 1021 EPLGSKTEFLDRAEMPTCPIONEPTRPENSYTPIKA1ENNGSDEFLCTEWKSN 1080
 Qy 1081 SVTSHQLRPADIKVVALGSLTAVGARNPNSDLPSTSRLGSLNISIGGNNLEHTT 1140
 Do 1081 SVTSHQLRPADIKVVALGSLTAVGARNPNSDLPSTSRLGSLNISIGGNNLEHTT 1140
 Qy 1141 LPNTLKKENPYLGFSTSTWEGTAGLNVAEGARDMDPAQDWLVERMKNSPDINKEKD 1200
 Do 1141 LPNLKKENPYLGFSTSTWEGTAGLNVAEGARDMDPAQDWLVERMKNSPDINKEKD 1200
 Qy 1201 WKVLTFLFGNDLHYCENPEAHLATIVQHQIQUALILSESLPRTAVVNVVMEASLY 1260
 Do 1201 WKVLTFLFGNDLHYCENPEAHLATIVQHQIQUALILSESLPRTAVVNVVMEASLY 1260
 Qy 1261 QGGGKCMALAAQNCNTCLRHQSSELAQKELKVNWLQHGCSFSYWHQYQREDIAVV 1320
 Do 1261 QGGGKCMALAAQNCNTCLRHQSSELAQKELKVNWLQHGCSFSYWHQYQREDIAVV 1320
 Qy 1321 VQPFONTLTPNLERGOTDLDTFSEDCFFHSORGHAETMAIWNMMLEPVGKRTTSNFT 1380
 Do 1321 VQPFONTLTPNLERGOTDLDTFSEDCFFHSORGHAETMAIWNMMLEPVGKRTTSNFT 1380
 Qy 1381 HSRAKLKCPSPSPSYLPSYRNSYLTQKOPDKL 1440
 Do 1381 HSRAKLKCPSPSPSYLPSYRNSYLTQKOPDKL 1440
 Qy 1441 RRGGRDPPEMSIRTVAL 1458
 Do 1441 RRGGRDPPEMSIRTVAL 1458

RESULT 2
 ABB09556
 ID ABB09556 standard; Protein; 1458 AA.
 AC ABB09556;

XX DT 21-OCT-2002 (first entry)
 XX DE Human lipase NHL (Val 1318 variant).
 XX KW Human; lipase; chromosome 2; phospholipase B homologue; thyroid; brain;
 KW inflammatory disease; proliferative disease; infectious disease;
 KW clotting disorder; cancer; drug screening; mental disorder; NHL;
 KW diagnostic reagent; clinical trial monitoring; cosmetic; nutriceutical;
 KW mutation detection; gene expression analysis; transgenic animal;
 KW nootropic; cytostatic; antiinflammatory; single nucleotide polymorphism;
 KW SNP; gene therapy; enzyme.
 XX OS Homo sapiens.
 XX FH key Location/Qualifiers
 FT misc-difference 1318 /note= "Ala replaces Val in a polymorphic variant
 (ABB09555)"
 FT XX WO20025932B-A1.
 XX PD 01-AUG-2002.
 XX PF 22-JAN-2002; 2002WO-US01715.
 PR 24-JAN-2001; 2001US-264049P.
 XX PA (LEXI-) LEXICON GENETICS INC.
 XX PT Yu X, Miranda M, Turner CA;
 DR N-PSDB; ABO77624.
 XX PT Polynucleotides encoding human lipases that are structurally related to
 PT animal lipases, particularly phospholipase B, useful for drug
 PT screening, diagnosis and in gene therapy of biological disorders -
 XX DS Disclosure; Page - ; 44pp; English.
 XX CC The invention relates to a novel human lipase (NHL; ABB09555; ABB09556);
 CC and to nucleic acids encoding it (AB07763, AB07764). The NHL has
 CC structural similarity with animal lipases, particularly phospholipase B.
 CC Polynucleotides encoding NHL were obtained using human genomic sequences
 CC in conjunction with human thyroid and brain cDNAs. The NHL gene is
 CC located on chromosome 2, and contains a C/T polymorphism at position
 CC 3953 of the open reading frame (ORF), resulting in an Ala/Val
 CC substitution at position 131 in the protein. NHL nucleotides and
 CC proteins are useful for treating disorders such as inflammatory or
 CC proliferative disease, infectious disease, clotting disorders, and
 CC cancer. They can also be used in screening for compounds useful in
 CC the treatment of mental, biological or medical disorders, as diagnostic
 CC reagents, in clinical trial monitoring and in cosmetic and nutraceutical
 CC applications. NHL nucleotides can additionally be used in the detection
 CC of disease-associated mutations, in the analysis of gene expression, for
 CC the recombinant expression of NHL, to generate transgenic animals, in
 CC gene therapy, and as part of ribozyme and/or triple helix sequences
 CC useful in the modulation of NHL gene expression. The present sequence
 CC represents the Val 1318 variant of NHL.
 CC Note: The present sequence is not shown in the specification, but
 CC was derived from the the information given on page 18 and the Ala 1318
 CC NHL variant (AB09555) given in the sequence listing.
 XX SQ Sequence 1458 AA;

Query Match 99.9%; Score 7762; DB 23; Length 1458;
 Best Local Similarity 99.9%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 1457; Conservative 1.

OY 1 MGLRPGIFPLELMLLLGQSTPQHTSRKSTTREGQMLPETLNNSPPFCNPNCQVNNPSK 60
 Db 1 MGURPGIFPLELMLLLGQSTPQHTSRKSTTREGQMLPETLNNSPPFCNPNCQVNNPSK 60

OY 61 SVMSLKKSDIKEVVAIGNEIPDPGPGDLEKDWTERRQVCMGVMTVLSDIRYSPS 120
 Db 61 SVMSLKKSDIKEVVAIGNEIPDPGPGDLEKDWTERRQVCMGVMTVLSDIRYSPS 120

OY 121 VPMVPCHGKRVIPHGDAGEDMIAQELVNRNKENLQDFDWKLINVFNSNASOCYL 180
 Db 121 VPMVPCHGKRVIPHGDAGEDMIAQELVNRNKENLQDFDWKLINVFNSNASOCYL 180

OY 181 PSQQNGLAAGGYDELAVGVLQDQEVPRAFNVLDSSEAVSVRSQHGTWSPAPCCN 240
 Db 181 PSQQNGLAAGGYDELAVGVLQDQEVPRAFNVLDSSEAVSVRSQHGTWSPAPCCN 240

OY 241 CSBETTRILAKVNMQWSYQAWNSLASSRYBQESFTVVFQDPFFYETPSLHSEDPRQD 300
 Db 241 CSBETTRILAKVNMQWSYQAWNSLASSRYBQESFTVVFQDPFFYETPSLHSEDPRQD 300

OY 301 STTLAWHLWNRMNPEAGEKDEPSLVKHGRPMKCPDQSPYLSYRNSYNTLQLKQDKL 360
 Db 301 STTLAWHLWNRMNPEAGEKDEPSLVKHGRPMKCPDQSPYLSYRNSYNTLQLKQDKL 360

OY 361 EYREGAEIRCPKQPSITVPTVSVRHLKADINVIGALDSLTAQNGQSTPSVNLDTQ 420
 Db 361 EYREGAEIRCPKQPSITVPTVSVRHLKADINVIGALDSLTAQNGQSTPSVNLDTQ 420

OY 421 YRGLSWSGGDENIGTVTLANTIREPNSLKXGSPVSGTGKTSPTNAFLQAVAGGRADL 480
 Db 421 YRGLSWSGGDENIGTVTLANTIREPNSLKXGSPVSGTGKTSPTNAFLQAVAGGRADL 480

OY 481 PYQARRVMDNDTRTHFOEDWKLITIFCNDLCIFCNDLVHYSQNFNTNIGALDI 540
 Db 481 PYQARRVMDNDTRTHFOEDWKLITIFCNDLCIFCNDLVHYSQNFNTNIGALDI 540

OY 541 IHAEVPRAFVNLVTLERVLNLRELYKORKVTCRPMILRLSCCVLKEDDNSTBLATLBF 600
 Db 541 IHAEVPRAFVNLVTLERVLNLRELYKORKVTCRPMILRLSCCVLKEDDNSTBLATLBF 600

OY 601 NKQKQEKTHQHOLTESGRYDVTREDFTVVQVQPFENVDMRRTSEGUDPNSSFPADCCHFSKS 660
 Db 601 NKQKQEKTHQHOLTESGRYDVTREDFTVVQVQPFENVDMRRTSEGUDPNSSFPADCCHFSKS 660

OY 661 HERRASALWNMLNLPYQCKTRHKPEKNINTCPNQWPQRTYKOSMIGQTWLCRDR 720
 Db 661 HERRASALWNMLNLPYQCKTRHKPEKNINTCPNQWPQRTYKOSMIGQTWLCRDR 720

OY 721 APSAHLHTSVHALPADIQVVALGSLTLAGNGIGSKPDPLDVTOYRGUSYAGGDS 780
 Db 721 APSAHLHTSVHALPADIQVVALGSLTLAGNGIGSKPDPLDVTOYRGUSYAGGDS 780

OY 781 LENVTUTLNTIREFRNLTGVAVGTDANDTNAFLQAVPGKAEDIMSQVOTLQKWD 840
 Db 781 LENVTUTLNTIREFRNLTGVAVGTDANDTNAFLQAVPGKAEDIMSQVOTLQKWD 840

OY 841 DHRVNFEEDMKVITVLGGSQDLYCDTDSNLYSAAFVDHNLALVHREPVULNV 900
 Db 841 DHRVNFEEDMKVITVLGGSQDLYCDTDSNLYSAAFVDHNLALVHREPVULNV 900

OY 901 DELNPITMRQVLGNPDKCQVQASLCLNCULTRENSQELARLEAFSAYRSSMBLVG 960
 Db 901 DELNPITMRQVLGNPDKCQVQASLCLNCULTRENSQELARLEAFSAYRSSMBLVG 960

OY 961 SGRYDQEDSFVLOPFFONTOLPVLAQDPLTSFAPPDCITHPNQKPHSOLARALMTML 1020
 Db 961 SGRYDQEDSFVLOPFFONTOLPVLAQDPLTSFAPPDCITHPNQKPHSOLARALMTML 1020

OY 1021 EPLGSKETEIDLAEMPTCTPONERLRTTRNSNTYPIKPALEWMSDELTENKSN 1080
 Db 1021 EPLGSKETEIDLAEMPTCTPONERLRTTRNSNTYPIKPALEWMSDELTENKSN 1080

OY 1081 SVPTSMQLRPADIKVALGDSLTTAVGARNPNSSDLPTSWRGWSWISGGDNLHETTT 1140
 Db 1081 SVPTSMQLRPADIKVALGDSLTTAVGARNPNSSDLPTSWRGWSWISGGDNLHETTT 1140

| | | | | |
|----------|--|---|------|---|
| QY | 1141 | LPEILKKNPYPLGPSTSTWEGTAGLNVAEAGARADNPAQAWDLVERMKNSPDINLEKD | 1200 | FT Active-site 394..404 /note= "Lipolytic enzyme G-D-S-L family |
| Db | 1141 | LPIILKKNPYPPLGPSTSTWEGTAGLNVAEAGARADNPAQAWDLVERMKNSPDINLEKD | 1200 | FT Domain 740..868 /note= "Serine active-site" |
| QY | 1201 | WLKVTLFLGIVNDLHYCENPEAHATEVWHTQOALDISELPRAFYNNVNEASLY | 1260 | FT Active-site /note= "Lipase/acylhydrolase with GDSL-like motif" |
| Db | 1201 | WLKVTLFLGIVNDLHYCENPEAHATEVWHTQOALDISELPRAFYNNVNEASLY | 1260 | FT Active-site /note= "Lipolytic enzyme G-D-S-L family |
| QY | 1261 | QSGGGKCAMLAQNCTCLRSQSSLERKELKVNWLQHGSFSYHOTREDFAV | 1320 | FT Domain 1096..1219 /note= "Lipase/acylhydrolase with GDSL-like motif" |
| Db | 1261 | QSGGGKCAMLAQNCTCLRSQSSLERKELKVNWLQHGSFSYHOTREDFAV | 1320 | FT Domain 1415..1442 /note= "Transmembrane domain" |
| QY | 1321 | VQPFQNTLTPNERGDTITFSECDCHPSRDRCHAEMIALNMMLEPVGKTSNFT | 1380 | FT Domain XX PN WO200294988-A2. |
| Db | 1321 | VQPFQNTLTPNERGDTITFSECDCHPSRDRCHAEMIALNMMLEPVGKTSNFT | 1380 | FT Domain XX PD 28-NOV-2002. |
| QY | 1381 | HSSRAKLKCPSPEPSYLYTURNSRLPDPDAAEAEPEVLYWAUPVAVGVLGVIGI | 1440 | FT Domain XX PR 17-MAY-2002; 2002WO-US15688. |
| Db | 1381 | HSSRAKLKCPSPEPSYLYTURNSRLPDPDAAEAEPEVLYWAUPVAVGVLGVIGI | 1440 | FT Domain XX PR 18-MAY-2001; 2001US-292242P. |
| QY | 1441 | RGERGRREOPPMRLTVAL 1458 | 1441 | FT Domain XX PR 25-MAY-2001; 2001US-293726P. |
| Db | 1441 | RGERGRREOPPMRLTVAL 1458 | 1441 | FT Domain XX PR 01-JUN-2001; 2001US-303402P. |
| QY | | | | PR 06-JUL-2001; 2001US-314754P. |
| Db | | | | PR 24-AUG-2001; 2001US-314754P. |
| | | | | PR 22-JAN-2002; 2002US-351262P. |
| | | | | PR 29-MAR-2002; 2002US-368799P. |
| RESULT 3 | | | | XX |
| ID | AAE34440 | AAE34440 Standard; Protein; 1458 AA. | | PA (INCY) INCYTE GENOMICS INC. |
| XX | | | | XX |
| AC | AAE34440: | | | PT Tang YT, Yue H, Azinai Y, Baughn MR, Burford N, Reddy R; |
| XX | | | | PT Walia NK, Das D, Nguyen DB, Yao MG, Arvizu CS, Lu Y, Gandhi AR; |
| DT | 14-MAY-2003 (first entry) | | | PT Griffin JA, Elliott VS, Ramkumar J, Lal PG, Lu DAM, Lee EA; |
| XX | | | | PT Lee SY, Yue H, Yang J, Tribouley CM, Kable AE, Swarnakar A; |
| DE | Human lipid-associated molecule (LIPAM)-1 protein. | | | XX |
| XX | | | | DR WPI; 2003-120797/11. |
| KW | Human; lipid-associated molecule; LIPAM; cardiovascular disorder; stroke; arteriovenous fistula; atherosclerosis; hypertension; Raynaud's disease; aneurysm; congestive heart failure; thromboembolitis; angina pectoris; ischaemic heart disease; rheumatic heart disease; peptic oesophagitis; gastrointestinal disorder; lipid metabolism disorder; Crohn's disease; nausea; peptic ulcer; fatty liver; Fabry's disease; Gaucher's disease; diabetes mellitus; hyperlipidaemia; hypercholesterolaemia; epilepsy; acquired immunodeficiency syndrome; anaemia; Alzheimer's disease; asthma; gene therapy; autoimmune disorder; inflammatory disorder; kuru; dementia; prion disease; Creutzfeldt-Jakob disease; leukemia; cancer; adenocarcinoma; lymphoma; melanoma; myeloma; sarcoma; gene therapy; protein replacement therapy. | | | DR N-PADB; AADB2626; |
| OS | Homo sapiens. | | | XX |
| XX | | | | PT New human lipid-associated molecule (LIPAM) proteins and polyribonucleotides, useful for diagnosing, treating or preventing cardiovascular disorders (e.g. atherosclerosis), neurological disorders (e.g. Parkinson's disease) or cancers - |
| FH | Key Location/Qualifiers | | | CC The present invention relates to novel human lipid-associated molecules (LIPAM) and polyribonucleotides encoding such proteins. Sequences of the invention are useful for treating diseases or conditions associated with decreased expression of functional LIPAM. The antagonist is useful for treating a disease or condition associated with the overexpression of functional LIPAM. They are useful for diagnosing, treating or preventing cardiovascular disorders (e.g. arteriovenous fistula, atherosclerosis, hypertension, Raynaud's disease, aneurysms, varicose veins, congestive heart failure, thromboembolitis, angina pectoris, ischaemic heart disease or rheumatic heart disease); gastrointestinal disorders (e.g. peptic oesophagitis, nausea, peptic ulcer or Crohn's disease), lipid metabolism disorders (e.g. fatty liver, Fabry's disease, Gaucher's disease, diabetes mellitus, hyperlipidaemia, hypercholesterolaemia), autoimmune disorders or inflammatory disorders (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's disease), neurological disorders (e.g. stroke, epilepsy, dementia, Alzheimer's disease, or prion diseases such as kuru or Creutzfeldt-Jakob disease) or cancers (e.g. adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma or sarcoma). They are also used in gene therapy and protein replacement therapy. The present sequence is human |
| FT | Peptide 1..19 | /label= Signal-peptide | | CC LIPAM-1 protein. |
| FT | Protein 20..1458 | /note= "Human mature LIPAM-1 protein" | | XX |
| FT | Peptide 1..21 | /label= Signal-peptide | | Sequence 1458 AA; |
| FT | Protein 22..1458 | /note= "Human mature LIPAM-1 protein" | | Query Match 99.9%; Score 7759; DB 24; Length 1458; Best Local Similarity 99.9%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 1457; Conservative 0; /note= "Human mature LIPAM-1 protein" |
| FT | Peptide 1..23 | /label= Signal-peptide | | 1 MGLRPGLITLLELLILLGCGTCPOHTSPRKSTLREGQLMWPETLKNSPFRCNPNLKGVNMSK 60 /note= "Lipase/acylhydrolase with GDSL-like motif" |
| FT | Protein 24..1458 | /note= "Human mature LIPAM-1 protein" | | |
| FT | Peptide 1..27 | /label= Signal-peptide | | |
| FT | Protein 28..1458 | /note= "Human mature LIPAM-1 protein" | | |
| FT | Domain 393..521 | /note= "Lipase/acylhydrolase with GDSL-like motif" | | |

| | | | | | |
|----|---|------|----|---|------|
| Db | 1 MGLRGPGIELLELULLLGGCTPQHTSPRKSTLSQMLPPCPNPKLGVNNPSK | 60 | QY | 1141 LPNLTKKENPKNYLIGFSTSTWEGTAGLNVAEAGARDNAQAMDLVERKNNSPDINERD | 1200 |
| QY | 61 SVHSLKPSDIKFAIGNLEIPPPGTDLEKDWTERPQQENGWMTVLSDITRYFPS | 120 | Db | 1141 LPNLKKENPKNYLIGFSTSTWEGTAGLNVAEAGARDNAQAMDLVERKNNSPDINERD | 1200 |
| Db | 61 SVHSLKPSDIKFAIGNLEIPPPGTDLEKDWTERPQQENGWMTVLSDITRYFPS | 120 | QY | 1201 WKLVTLFFGVNDLHYCNCPEAHATEVQHQALDILSEELPRAFYNNVWELASSLY | 1260 |
| QY | 121 VPMFVCHTGKRVTPHDGAEDLWIAQELVRNWKENLQDFDOWKLINUFPFSVNASCYLC | 180 | Db | 1201 WKLVTLFFGVNDLHYCNCPEAHATEVQHQALDILSEELPRAFYNNVWELASSLY | 1260 |
| Db | 121 VPMFVCHTGKRVTPHDGAEDLWIAQELVRNWKENLQDFDOWKLINUFPFSVNASCYLC | 180 | QY | 1261 QGGGKCMALAQNCTCLRHSSOLEKEUKUNNIQHGSSPSYTHQYQREDFAW | 1320 |
| QY | 181 PSAQONGLAAGGDELGMGVYDYLQOEVPRAFVNLVDJSEBAEVNSQYGTWISAPERCN | 240 | Db | 1261 QGGGKCMALAQNCTCLRHSSOLEKEUKUNNIQHGSSPSYTHQYQREDFAW | 1320 |
| Db | 181 PSAQONGLAAGGDELGMGVYDYLQOEVPRAFVNLVDJSEBAEVNSQYGTWISAPERCN | 240 | QY | 1321 VOPPFONTLTPLNERGDTDLTPSEDCFHSRGAEMAIALWNMPLPVGRKITSNIFT | 1380 |
| QY | 241 CSBETTRILAKVUMQWSOBANSLASRSYSEBSFTVFORFYETPLSISSEDPRLQD | 300 | Db | 1321 VOPPFONTLTPLNERGDTDLTPSEDCFHSRGAEMAIALWNMPLPVGRKITSNIFT | 1380 |
| Db | 241 CSBETTRILAKVUMQWSOBANSLASRSYSEBSFTVFORFYETPLSISSEDPRLQD | 300 | QY | 1331 HSPAKLKCPSPSPSYLVTNRLLPQAEAPVLTWAVPAAGVGWVWGLITVWR | 1440 |
| QY | 301 STTLAWHLWNRMMEPAGEKDEPLSVKGRPMKPSQESPYLSYRNSTLTHQKPDQ | 360 | Db | 1331 HSPAKLKCPSPSPSYLVTNRLLPQAEAPVLTWAVPAAGVGWVWGLITVWR | 1440 |
| Db | 301 STTLAWHLWNRMMEPAGEKDEPLSVKGRPMKPSQESPYLSYRNSTLTHQKPDQ | 360 | QY | 1441 RRGRRREPPMSLRTVAL | 1458 |
| QY | 361 EYREGAETROPDKDPSDVTPTSHRLKADINVAGLDSTAGNGAGSTPSPVNLVDITQ | 420 | Db | 1441 RRGRRREPPMSLRTVAL | 1458 |
| Db | 361 EYREGAETROPDKDPSDVTPTSHRLKADINVAGLDSTAGNGAGSTPSPVNLVDITQ | 420 | QY | 421 YRGLSWSYGGDENITGTVTLANTIREFPLSKFSGVCKETSPNAFINAQAGGRAELD | 480 |
| QY | 421 YRGLSWSYGGDENITGTVTLANTIREFPLSKFSGVCKETSPNAFINAQAGGRAELD | 480 | Db | 480 YQARREYDLMKDTRIFQEDMKLITIFIGANDLCPDFNDLHYSPONFTNIGKLDI | 540 |
| Db | 480 YQARREYDLMKDTRIFQEDMKLITIFIGANDLCPDFNDLHYSPONFTNIGKLDI | 540 | QY | 481 PVQARRLYDLMKDTRIFQEDMKLITIFIGANDLCPDFNDLHYSPONFTNIGKLDI | 540 |
| QY | 541 LHAEVPRAFVNLTVEUNRLIQEKKVYCPMLASLCPVLUKFEDNSBLATIEF | 600 | Db | 541 LHAEVPRAFVNLTVEUNRLIQEKKVYCPMLASLCPVLUKFEDNSBLATIEF | 600 |
| Db | 541 LHAEVPRAFVNLTVEUNRLIQEKKVYCPMLASLCPVLUKFEDNSBLATIEF | 600 | QY | 601 NKKFOEKTHOLIESGRYPTREDFVVVQPFENMDPKTSBGPDPNSFPAPPCHFSSKS | 660 |
| QY | 601 NKKFOEKTHOLIESGRYPTREDFVVVQPFENMDPKTSBGPDPNSFPAPPCHFSSKS | 660 | Db | 601 NKKFOEKTHOLIESGRYPTREDFVVVQPFENMDPKTSBGPDPNSFPAPPCHFSSKS | 660 |
| QY | 661 HSRASAASLWNMMLEPVGOKTTRHKPFENKINTCPNQVOPFLRTYKNSQHGTWLRCDR | 720 | Db | 661 HSRASAASLWNMMLEPVGOKTTRHKPFENKINTCPNQVOPFLRTYKNSQHGTWLRCDR | 720 |
| Db | 661 HSRASAASLWNMMLEPVGOKTTRHKPFENKINTCPNQVOPFLRTYKNSQHGTWLRCDR | 720 | QY | 721 APLALHPTSVHAIHPADQVQWAAGDSLTTAGNGTSKDDLPDTTOYRGLSYAGGDS | 780 |
| QY | 721 APLALHPTSVHAIHPADQVQWAAGDSLTTAGNGTSKDDLPDTTOYRGLSYAGGDS | 780 | Db | 721 APLALHPTSVHAIHPADQVQWAAGDSLTTAGNGTSKDDLPDTTOYRGLSYAGGDS | 780 |
| Db | 721 APLALHPTSVHAIHPADQVQWAAGDSLTTAGNGTSKDDLPDTTOYRGLSYAGGDS | 780 | QY | 781 LENVTTLNIRLEPFRNRLTYGAVGTGDAINTAFLNQAVPGAKADEJMSQVTLMQKRD | 840 |
| Db | 781 LENVTTLNIRLEPFRNRLTYGAVGTGDAINTAFLNQAVPGAKADEJMSQVTLMQKRD | 840 | Db | 781 LENVTTLNIRLEPFRNRLTYGAVGTGDAINTAFLNQAVPGAKADEJMSQVTLMQKRD | 840 |
| QY | 841 DHRVNPHEDWKVITVLIGSDCDDCYTSNLSAANFDHJRALDVJHREPPVNUV | 900 | Db | 841 DHRVNPHEDWKVITVLIGSDCDDCYTSNLSAANFDHJRALDVJHREPPVNUV | 900 |
| Db | 841 DHRVNPHEDWKVITVLIGSDCDDCYTSNLSAANFDHJRALDVJHREPPVNUV | 900 | QY | 901 DFLNPTIMRQVFLGNPDKCPVQOASVLNCVLTURENSQBABLEAFSRAYSSMRELV | 960 |
| QY | 901 DFLNPTIMRQVFLGNPDKCPVQOASVLNCVLTURENSQBABLEAFSRAYSSMRELV | 960 | Db | 901 DFLNPTIMRQVFLGNPDKCPVQOASVLNCVLTURENSQBABLEAFSRAYSSMRELV | 960 |
| Db | 901 DFLNPTIMRQVFLGNPDKCPVQOASVLNCVLTURENSQBABLEAFSRAYSSMRELV | 960 | QY | 961 SGYDQTOEDFSVVLQPFONIOLPVLADGLPDSFFADCPCHINKPNQKHSQALARWNTML | 1020 |
| QY | 961 SGYDQTOEDFSVVLQPFONIOLPVLADGLPDSFFADCPCHINKPNQKHSQALARWNTML | 1020 | Db | 961 SGYDQTOEDFSVVLQPFONIOLPVLADGLPDSFFADCPCHINKPNQKHSQALARWNTML | 1020 |
| QY | 1081 SVTSTVHQLRPADIKVVALGDSITTAVGARPNNSDILPTSRGLSNSIGGONLETT | 1140 | Db | 1081 SVTSTVHQLRPADIKVVALGDSITTAVGARPNNSDILPTSRGLSNSIGGONLETT | 1140 |
| Db | 1081 SVTSTVHQLRPADIKVVALGDSITTAVGARPNNSDILPTSRGLSNSIGGONLETT | 1140 | QY | 1141 LPNLTKKENPKNYLIGFSTSTWEGTAGLNVAEAGARDNAQAMDLVERKNNSPDINERD | 1200 |
| Db | 1141 LPNLKKENPKNYLIGFSTSTWEGTAGLNVAEAGARDNAQAMDLVERKNNSPDINERD | 1200 | Db | 1141 LPNLKKENPKNYLIGFSTSTWEGTAGLNVAEAGARDNAQAMDLVERKNNSPDINERD | 1200 |
| PR | 30-MAR-2001; 2001US-279959P. | | PR | 26-MAR-2001; 2001US-277321P. | |
| PR | 27-MAR-2001; 2001US-278999P. | | PR | 20-MAR-2001; 2001US-277327P. | |
| PR | 28-MAR-2001; 2001US-279444P. | | PR | 21-MAR-2001; 2001US-277944P. | |
| PR | 30-MAR-2001; 2001US-27738P. | | PR | 22-MAR-2001; 2001US-27783P. | |
| PR | 30-MAR-2001; 2001US-279959P. | | PR | 23-MAR-2001; 2001US-278132P. | |
| PR | 30-MAR-2001; 2001US-280233P. | | PR | 26-MAR-2001; 2001US-277944P. | |

| | | |
|----------|------|--|
| SQ | PR | 02-APR-2001; 2001US-280802P. |
| | PR | 02-APR-2001; 2001US-280802P. |
| | PR | 04-APR-2001; 2001US-281194P. |
| | PR | 13-APR-2001; 2001US-283675P. |
| | PR | 30-APR-2001; 2001US-287424P. |
| | PR | 02-MAY-2001; 2001US-288066P. |
| | PR | 03-MAY-2001; 2001US-288342P. |
| | PR | 03-MAY-2001; 2001US-288528P. |
| | PR | 15-MAY-2001; 2001US-291190P. |
| | PR | 16-MAY-2001; 2001US-291199P. |
| | PR | 16-MAY-2001; 2001US-291240P. |
| | PR | 30-MAY-2001; 2001US-294485P. |
| | PR | 31-JULY-2001; 2001US-294889P. |
| | PR | 16-AUG-2001; 2001US-294999P. |
| | PR | 18-JUN-2001; 2001US-299027P. |
| | PR | 19-JUN-2001; 2001US-299303P. |
| | PR | 10-JULY-2001; 2001US-294354P. |
| | PR | 31-JULY-2001; 2001US-309198P. |
| | PR | 10-SEP-2001; 2001US-318462P. |
| | PR | 12-SEP-2001; 2001US-31870P. |
| | PR | 27-SEP-2001; 2001US-325430P. |
| | PR | 18-OCT-2001; 2001US-325681P. |
| | PR | 31-OCT-2001; 2001US-335301P. |
| | PR | 14-NOV-2001; 2001US-332172P. |
| | PR | 14-NOV-2001; 2001US-332272P. |
| | PR | 14-NOV-2001; 2001US-333184P. |
| | PR | 14-NOV-2001; 2001US-333272P. |
| | PR | 21-NOV-2001; 2001US-332094P. |
| | PR | 03-DEC-2001; 2001US-337426P. |
| | PR | 03-DEC-2001; 2001US-338092P. |
| | PR | 04-DEC-2001; 2001US-337185P. |
| | PR | 03-JAN-2002; 2002US-345705P. |
| | PR | 07-MAR-2002; 2002US-0092900. |
| | PA | (CURA-) CURAGEN CORP. |
| | XX | Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CER, Li L; Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R; Parihar A, Ganguli E, Vermet CAM, Guo X, Tchernev V; Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y; Anderson D, Spaderna SK, Catterton B, Burgess C, Leite M, Zhong H; Alisbrook JP, Lepley DM, Rieger DK; |
| | PR | WPI; 2002-723337/8. |
| | DR | N-PSDB; ABX97050. |
| | XX | PT NOKV polypeptides and poly nucleotides, useful for preventing or treating a disorder associated with aberrant NOX expression or activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial asthma - |
| | PS | Claim 1; Page 170; 1103pp; English. |
| | PT | This invention describes novel human NOX polypeptides which have cytostatic, cardiotonic, antiarrtherosclerotic, antiasthmatic and hypotensive activity. Pharmaceutical compositions comprising the NOX proteins or nucleic acid molecules or NOX antibodies are useful for preventing or treating a disorder associated with aberrant NOX expression or activity e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial asthma. The products of the invention can be used for gene therapy or in a vaccine. ABU65041-ABU65218 represent the NOX polypeptides encoded by ABX97008-ABX97185. |
| Sequence | QQ | 1423 AA; |
| | 1 | MGLRPGLFLBLLLLGQSTPPQINTSPRKSTLQGQMLPTRLKNSPPPCPNPKLGUNMPSK |
| | 1 | MGLRPGLFLBLLLLGQCTPQINTSPRKSTLQGQMLPTRLKNSPPPCPNPKLGUNMPSK |
| | 40 | |
| | 61 | SVHSLKPSDIKFAIGNILEIPPPGPGTQDLEKQDWTERPQQCNGVMTVLSDIRYFSPS |
| | 120 | |
| | 41 | -VHSLKPSDIKFAIGNILEIPPPGPGTQDLEKQD--ERPQQCNGVMTVLSDIRYFSPS |
| | 97 | |
| | 121 | VPMVCHGCKRVPHGDAELWQAEULTRANKMENQLDQFDWKLNIVFFSNASOYLC |
| | 180 | |
| | 98 | VPMVCHGCKRVPHGDAELWQAEULTRANKMENQLDQFDWKLNIVFFSNASOYLC |
| | 156 | |
| | 181 | PSAQONGLAGGIVDELMGFLDYLQOEVRPAFLVNLDSVAEVSRQHGTWLSAPECN |
| | 240 | |
| | 217 | CSEETRLAKVMQWSYQBAWNSLASSRSYEQSSTVVFQPFYETTS---DPLQD |
| | 272 | |
| | 301 | STTLAWHLNNRMMEPAGEKDEPLSVKHPGMKPCPSQESTYFLSTRNSNTLTQDKPL |
| | 360 | |
| | 273 | STTLAWHLNNRMMEPAGEKDEPLSVKHPGMKPCPSQESTYFLSTRNSNTLTQDKPL |
| | 332 | |
| | 361 | EYREGAETRCPDKPSDTPTSVHLKPADINVAGLGSUTAGNGAGSPPGNYLVDLQ |
| | 420 | |
| | 333 | EYREGAETRCPDKPSDTPTSVHLKPADINVAGLGSUTAGNGAGSPPGNYLVDLQ |
| | 392 | |
| | 421 | YRGLSWVGEGDENIGTVTLIAKQDFNLSLKGSGVGTGKETSNAFLNQAVAGRAEDL |
| | 480 | |
| | 393 | YRGLSWVGEGDENIGTVTLIAKQDFNLSLKGSGVGTGKETSNAFLNQAVAGRAE-- |
| | 450 | |
| | 481 | PVQARLVDLMDNQTRTHQBDWKHITLFGNDICDFCMDLVMSQPOHTDNIGALD |
| | 540 | |
| | 451 | --QARRLVLDIMKDNTRIHQBDWKHITLFIGNDICDFCMDLVMSQPOHTDNIGALD |
| | 508 | |
| | 541 | LHAEVPRAFTVNLVTLVNLRELYQEKVYCPMILSLCPCTLKDENDSTELATLIEF |
| | 600 | |
| | 509 | LHAEVPRAFTVNLVTLVNLRELYQEKVYCPMILSLCPCTLKDENDSTELATLIEF |
| | 568 | |
| | 601 | NKKQEKTHOLIESCGYDTRBDFWVQPFENUDMPKTSQEGUDNSPAPDCPHFSKS |
| | 660 | |
| | 569 | NKKQEKTHOLIESCGYDTRBDFWVQPFENUDMPKTSQEGUDNSPAPDCPHFSKS |
| | 628 | |
| | 661 | HISRAASALAMNMLBPVGQKTRHAFENKINTCNOVO_PFLRTYKNSQHGHWLPCRD |
| | 719 | |
| | 629 | HISRAASALAMNMLBPVGQKTRHAFENKINTCNOVO_PFLRTYKNSQHGHWLPCRD |
| | 688 | |
| | 720 | RAPSALHPTSVHLRPAQDIAVAGLSDLTAGNCIGSKDDLPVTTOYRGLSYAGGD |
| | 779 | |
| | 689 | RAPSALHPTSVHLRPAQDIAVAGLSDLTAGNCIGSKDDLPVTTOYRGLSYAGGD |
| | 748 | |
| | 780 | SLENTVTNLIREFRNRTGAYGTGQDANTTMFLNQAVGAGCAGDMQSQVQTMQMK |
| | 839 | |
| | 749 | SLENTVTNLIREFRNRTGAYGTGQDANTTMFLNQAVGAGCAGDMQSQVQTMQMK |
| | 808 | |
| | 840 | DDHRIINFHDKVIVLIGSDDCYCDSMISAANFDHLRNLVJREVERVLN |
| | 899 | |
| | 809 | DHHRNFHDKVIVLIGSDDCYCDSMISAANFDHLRNLVJREVERVLN |
| | 868 | |
| | 900 | VDFLNPTIMQVFLGPNPKCWPQOASVLCNCVLTURENQSOELARLEAFSRAYRSMRELV |
| | 959 | |
| | 869 | WDFLNPTIMQVFLGPNPKCWPQOASVLCNCVLTURENQSOELARLEAFSRAYRSMRELV |
| | 928 | |
| | 960 | GSGRYDQTDPFSVLUQPFFONIOLPVLAQGDPDFFADCPICHNQKFISOLARALWTM |
| | 1019 | |
| | 929 | GSGRYDQTDPFSVLUQPFFONIOLPVLAQGDPDFFADCPICHNQKFISOLARALWTM |
| | 988 | |

| | | | |
|-----------------|--|----|---|
| QY | 1050 NSVPTSVHOLRADIKVAALGDSLTTAVGARNNSDLPTSGRLWSIGDGNEHT | PR | 30-MAR-2001; 2001US-277338P. |
| Db | 1049 NSVPTSVHOLRADIKVAALGDSLTTAVGARNNSDLPTSGRLWSIGDGNEHT | PR | 30-MAR-2001; 2001US-279958P. |
| QY | 1140 TLPNLKFPNPLGLGFSSTTWEGTAGLNVAAGARARDMPAQAWLVERMKSP-DINLE | PR | 02-APR-2001; 2001US-280233P. |
| Db | 1109 TLDILKFPNPLGLGFSSTTWEGTAGLNVAAGARARDMPAQAWLVERMKSP-DINLE | PR | 02-APR-2001; 2001US-280822P. |
| QY | 1199 KWKLVLTFIGNDLHYCENPEAHATEYVOMIQADLISBELPRAFVNVEMLAS | PR | 02-MAY-2001; 2001US-288066P. |
| Db | 1169 KWKLVLVILFIGNDLHYCENP---VGEIVHQIQLDILSBELPRAFVNVEMLAS | PR | 03-MAY-2001; 2001US-288528P. |
| QY | 1259 LYGOGSKCAMAQQNCTCLHSQSSEKELUKKWNLOGISSYWMQTYQREDFA | PR | 04-APR-2001; 2001US-281194P. |
| Db | 1225 LYOGOGSKCAMAQQNCTCLHSQSSEKELUKKWNLOGISSYWMQTYQREDFA | PR | 13-APR-2001; 2001US-283675P. |
| QY | 1319 VVQPFONTLPLNERGDTLTPPSDFHSDRGHMA-TALWNMLEPGRKTSNN | PR | 30-MAY-2001; 2001US-291240P. |
| Db | 1285 VVQPFONTLPLNERGDTLTPPSDFHSDRGHMA-TALWNMLEPGRKTSNN | PR | 02-MAY-2001; 2001US-287424P. |
| QY | 1379 FTHSRAKLKCPSPESPVYLTNRSLPDAEPAPEVLYWVPAAGVGLVVGIGTVWW | PR | 03-MAY-2001; 2001US-288342P. |
| Db | 1344 FTHSRAKLKCPSPESPVYLTNRSLPDAEPAPEVLYWVPAAGVGLVVGIGTVWW | PR | 03-MAY-2001; 2001US-294889P. |
| Qy | 1439 RRRGGREDPPLSLRTVAL 1458 | PR | 15-MAY-2001; 2001US-291190P. |
| Db | 1404 RRRGGREDPPLSLRTVAL 1423 | PR | 16-MAY-2001; 2001US-291095P. |
| RESULT 5 | | | |
| ABU65081 | ID ABU65081 standard; Protein: 1419 AA. | PR | 16-MAY-2001; 2001US-291095P. |
| XX | AC ABU65081; | PR | 16-MAY-2001; 2001US-291240P. |
| XX | DT 20-MAY-2003 (first entry) | PR | 30-MAY-2001; 2001US-294889P. |
| XX | DT 20-MAY-2003 (first entry) | PR | 31-MAY-2001; 2001US-294899P. |
| DE | Human NOV24a protein. | PR | 31-MAY-2001; 2001US-309188P. |
| XX | KW NOVX: cytotoxic; cardiant; antiarteriosclerotic; antiasthmatic; cancer; hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine; human. | PR | 18-JUN-2001; 2001US-299077P. |
| OS | Homo sapiens. | PR | 19-JUN-2001; 2001US-299303P. |
| XX | WO200372757-A2. | PR | 10-JUL-2001; 2001US-304548P. |
| PD | 19-SEP-2002. | PR | 31-JUL-2001; 2001US-309188P. |
| XX | PP 08-MAR-2002; 2002WO-US06908. | PR | 10-OCT-2001; 2001US-332054P. |
| PR | 08-MAR-2002; 2002WO-US06908. | PR | 11-OCT-2001; 2001US-335301P. |
| PR | 08-MAR-2002; 2002WO-US06908. | PR | 14-NOV-2001; 2001US-332172P. |
| PR | 08-MAR-2002; 2002WO-US06908. | PR | 14-NOV-2001; 2001US-332271P. |
| PR | 08-MAR-2002; 2002WO-US06908. | PR | 14-NOV-2001; 2001US-332272P. |
| PR | 08-MAR-2002; 2002WO-US06908. | PR | 14-NOV-2001; 2001US-333194P. |
| PR | 08-MAR-2002; 2002WO-US06908. | PR | 14-NOV-2001; 2001US-333272P. |
| PR | 08-MAR-2002; 2002WO-US06908. | PR | 21-NOV-2001; 2001US-332054P. |
| PR | 08-MAR-2002; 2002WO-US06908. | PR | 03-DEC-2001; 2001US-337426P. |
| PR | 08-MAR-2002; 2002WO-US06908. | PR | 03-DEC-2001; 2001US-338022P. |
| PR | 08-MAR-2002; 2002WO-US06908. | PR | 04-DEC-2001; 2001US-337185P. |
| PR | 08-MAR-2002; 2002WO-US06908. | PR | 03-JAN-2002; 2002US-345705P. |
| PR | 08-MAR-2002; 2002WO-US06908. | PR | 07-MAR-2002; 2002US-0092900. |
| PA | (CURA-) CURAGEN CORP. | | |
| XX | Padigaru M, Szymek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L; Zerhouni BD, Gubov V, Ji W, Gorman L, Miller CE, Kekuda R; Patterson M, Ganguli E, Varnet CAM, Guo X, Tchernev V; Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y; Anderson D, Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H; Alsobrook JP, Lepley DM, Rieger DK; DR N-PSDB; ABX97048. | FT | This invention describes novel human NOVX polypeptides which have cytotoxic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive activity. Pharmaceutical compositions comprising the NOVX proteins or nucleic acid molecules or NOVX antibodies are useful for preventing or treating a disorder associated with aberrant NOVX expression or activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial asthma. |
| XX | Claim 1, Page 168-169; 1103pp; English. | FT | This invention describes novel human NOVX polypeptides which have cytotoxic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive activity. Pharmaceutical compositions comprising the NOVX proteins or nucleic acid molecules or NOVX antibodies are useful for preventing or treating a disorder associated with aberrant NOVX expression or activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial asthma. The products of the invention can be used for gene therapy or in a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by ABX97008-ABX97185. |
| SQ | Sequence 1419 AA; | XX | sequence 1419 AA; |

| Query | Match | Score | DB | Length |
|-------|--|--|----------------------|---|
| QY | Best Local Similarity 84.2%; Pred. No. 0; Matches 1236; Conservative 57; Mismatches 100; Indels 75; Gaps 21; | 81.3%; Score 6311.5; DB 23; Length 1419; | Db | 978 ARAIWNTNLEPLASKTENLDRAREMPITCPTONEPFLTRPRNSNTYPIKPALENGSDF 1037 |
| QY | 7 IFLFLGLLMLGQTPQHPSRKRSTLKGQWPELTQNSPPCPNPKLGVNMP--SKSVH 56 | 11 VFLIGLILPFLGANCILQTSKGKCTLRGRY---PQPQP-----PCLSPVH 55 | QY | 1072 LCTEWKASSNSVPSVHQIARADIKVVALGLDSLTAVGARPNNSDFTSMRGLSMSIGG 1131 |
| QY | 64 SLPSDIKEVAIGNLETTPPDGTGDLKQD--WTERPQVQMGWMTLSPIRYSPS 120 | 115 RSLSLRREAVLPAAAGBESLCHDPLFQSGAGOLSPPPROWS-WFOQACLRPGVKKEMODVGRPS 114 | Db | 1038 LCTEWKASSNSVPSVHQIARADIKVVALGLDSLTAVGARPNNSDFTSMRGLSMSIGG 1097 |
| QY | 56 QLRPADIKVVALGLDSLTDFOBGSAGOLSPPPROWS-WFOQACLRPGVKKEMODVGRPS 114 | 121 VPMIVUHTCKRVRPHDGAE---DLWIOAQELVRNKENLQ-LDFQDWKLINVFFSNA 174 | QY | 1132 DGNLETHTTLPNLTKKENPYPYLGFSTTWEGTAGLNVAEGARA-RDMPAQAWDLVERMK 1190 |
| QY | 115 RSLSLRREAVLPAAAGBESLCHDPLFQSGAGOLSPPPROWS-WFOQACLRPGVKKEMODVGRPS 114 | 115 RSLSLRREAVLPAAAGBESLCHDPLFQSGAGOLSPPPROWS-WFOQACLRPGVKKEMODVGRPS 114 | Db | 1098 DGNELETHTTLPSLKKEPNPYLGFSTSWEGTAGLNVAEGARARRODMPAQAWDLVERMK 1157 |
| QY | 175 SQCYLCPSAQNGLAAGGVDELMQVGDYVQOQVYDQVNLVLDSEVAEVSRQHGTWLSP 234 | 174 DLCHTCYPLVQGPVFLDGLMDTILHSI----QFRAFNVVEMBLA---IQG---- 219 | QY | 1191 NSPDINIEKDWKVLTFLGNDLICHYENPEAHATEWQHQIQLQDLSLSELPRATVNV 1250 |
| QY | 235 APEPCNCSEBTETRALKVUNQMSQEAMSLASSRYSQESTFWFOPPFYRTPSLHE 294 | 220 -OOGKCA-----MLAQEAWNSILASSRYSQESTFWFOPPFYRTPS---- 262 | QY | 1158 NSP-IRHOEDWKITLTFIGGNDLDFCN-----LVGEVQHOIQLDLSLSELPRATVNV 1212 |
| Db | 295 DPLQDSTTLAHLWMARMMEPAGEKDEPLSVKGRMPCMKPSQESPYLSYRNNSYLRLQ 354 | 263 DPLQDSTTLAHLWNRMMEPAGEKDEPLSVKGRMPCMKPSQESPYLSYRNNSYLRLQ 322 | Db | 1251 VEMELASLYQOGGKCMALAAQNCTCLRHOSSELEKKVUNLQHGSFSFWHQ 1310 |
| QY | 355 KPDQKLEYREGAEITRCPDKDPDSDTVPVTSHRLKPADINVIGALGDSLTAGNAGGSTPNV 414 | 323 KPDQKL-WREGAEITRCPDKDPDSDTVPVSFLRKPADINVIGALGDSLTAGNAGGSTPNV 381 | QY | 1223 VEMLASLYQOGGKAMLAQONCTCLRHOSSELEKKVUNLQHGSFSFWHQ 1272 |
| QY | 415 LDVLTQYRGSLWSYGGDENIGTVTLANTIRENPSPKGSIGTKETSPTNATLNQAVG 474 | 382 LDVLTQYRGSLWSYGGDENIGTVTLADIRENPKLKFPSVGTGKETSPNAPLNQAVG 441 | QY | 1311 YTQREDFAVVVOPPFQNTLTPANERGDTDLTFSEDCFHSDRGHAENAIALWNMFLPV 1370 |
| Db | 442 GRAS----QARRVLDLMDKTRTHFQDWTITLFIGENDLCPFDNST 534 | 475 GRABDLVQARRLVUDLMDKTRHFOEDWKTITLFIGENDLCPFDNST 534 | Db | 1273 YTQREDFAVVVOPPFQNTLTPIN-RGDTDLTFSEDCFHSDRGHAENAIALWNMFLPV 1331 |
| QY | 535 KALDILHAE--VRAFAVNLVLTIEVNLRELTOEKKYCPRMTLRSICPCVFKFDNST 592 | 442 GRAS----QARRVLDLMDKTRTHFQDWTITLFIGENDLCPFDNST 497 | QY | 1371 GRKTSNNPHTSRAKLKCPSPESYLTLRNLRRPQAAEPREVLYWAVPVAAGVIVV 1430 |
| Db | 498 KALDILHAEQVRAFAVNLVLTIEVNLRELTOEKKYCPRMTLRSICPCVFKFDNST 557 | 593 EATLIEFEKKFOBKTHOLIESERYDTRBDTUVQPFENVKMPKTSSEGRLNSFFARD 652 | Db | 1332 GRKTSNNPHTSRAKLKCPSPESYLTLRNLRRPQAAEPREVLYWAVPVAAGVIVV 1391 |
| QY | 593 EATLIEFEKKFOBKTHOLIESERYDTRBDTUVQPFENVKMPKTSSEGRLNSFFARD 652 | 558 EATLIEFNKKFOBKTHOLIESGRYDTRBDTUVQPFENVKMPKTBGLDNSSFFAD 617 | QY | 1431 GIIGTVVWRCCRGRREPPMSLRTVAL 1458 |
| Db | 653 CFHSSKSRSRASALWNMPLVYQKOTTRHKPEKNKINTCPQVQPLRTYKNSMOSH 712 | 618 CFHSSKSRSRASALWNMPLVYQKOTTRHKPEKNKINTCPQVQPLRTYKNSMOSH 677 | Db | 1392 GIIGTVVWRCCRGRREPPMSLRTVAL 1419 |
| QY | 713 TWLPCRDRAPSALHPTSVHRLPADIQVVAALGLSLTAGNGLGSKPDLLDPTMORYRSL 772 | 678 TWLPCRDRAPSALHPTSVHRLPADIQVVAALGLSLTAGNGLGSKPDLLDPTMORYRSL 737 | QY | RESULT 6 |
| QY | 773 YSAAGGDSLLENVTLPSSILREFRNRLTYGAVGTDGANDTNAFLNQAVPGKAQDMSQV 831 | 738 YSAAGGDSLLENVTLPSSILREFRNRLTYGAVGTDGANDTNAFLNQAVPGKAQDMSQV 797 | QY | AAB22860 |
| Db | 892 VPRVLVNLVDPLNPTIMQVFLGNDPDKCEVQQAISVLCNCVLTRENSDLEARFSAY 951 | 832 QTLMOKMKDHRYNFHEDMKTIVTLLIGSDDCYCTDSNLYSANFHDLRNALDVRE 891 | Db | ID AAB22860 standard; Protein: 1216 AA. |
| QY | 858 VPRVLVNLVDPLNPTIMQVFLGNDPDKCEVQQAISVLCNCVLTRENSDLEARFSAY 951 | 798 QTMOKMKDHRYNFHEDMKTIVTLLIGSDDCYCTDSNLYSANFHDLRNALDVRE 857 | QY | XX AAB22860; |
| Db | 952 RSSRRELVGSGRYPTOEDSVLQPPFONIOLPVLADGLPDTSFAPDCIHPNOKFHSOL 1011 | 918 OSSRRELVGSGRYPTOEDSVLQPPFONIOLPVLQDGLPDTSFAPDCIHPNOKFHSOL 977 | Db | XX 10-OCT-2000; 2000US-238445P. |
| QY | 918 OSSRRELVGSGRYPTOEDSVLQPPFONIOLPVLQDGLPDTSFAPDCIHPNOKFHSOL 977 | 1012 ARALWNTNLEPLPGSKTETIDRAEMPITCPTQEPFLRTPRNSNTYPIKATENWGSDP 1071 | QY | XX 26-DEC-2000; 2000US-257293P. |
| Db | | | PA (FARB) BAYER AG. | |
| QY | | | PI Zhu Z; | |
| | | | XX | |

DR WPI; 2002-226287/45.
DR N-PSDB; AAD37410.

XX

PT New human phospholipase-like enzyme polypeptide useful for screening

PT agents, and in the treatment of cancer, inflammation, diabetes,

PT obesity, a central nervous system disorder, or a cardiovascular

PT disorder -

XX

PS Claim 25; Fig 2; 144pp; English.

CC The present invention relates to novel human phospholipase-like enzymes
CC and polynucleotides encoding such proteins. Sequences of the invention
CC are useful for producing a medicament for modulating the activity of
CC phospholipase in a disease such as cancer, inflammation, diabetes,
CC disorders, chronic obstructive pulmonary diseases, central nervous system
(CNS) disorder, such as brain injuries, cerebrovascular disease, dementia
CC (Alzheimer's disease), Parkinson's disease, Huntington's disease, corticobasal degeneration,
CC motor neuron disease, Pick's disease, Creutzfeld-
CC Jacob dementia, schizophrenia with dementia, Korsakoff's psychosis,
CC pain associated with CNS (e.g., epilepsy, failed back surgery syndrome,
CC sciatica), multiple sclerosis, stroke, age associated memory impairment,
CC allergic disease including asthma, allergic rhinitis (hay fever), atopic
CC dermatitis, anaphylaxis and inflammation, cardiovascular disease, chronic
CC obstructive pulmonary disease, acute respiratory distress syndrome, gout,
CC diabetes, emphysema or obesity. They are also used for treating anorexia,
CC overweight, cachexia, bulimia, hypertension, type-II diabetes, coronary,
CC artery disease, hyperlipidemia, gall bladder disease, osteoarthritis,
CC sleep apnoea and respiratory problems, cancer (e.g. breast, prostate,
CC ovarian cancer), thrombolytic disease, reduced fertility, polycystic
CC hirsutism, stress incontinence and depression. The present sequence
CC is human phospholipase like enzyme.

XX

SQ Sequence 1216 AA;

Query Match 80.5%; Score 6253.5; DB 23; Length 1216;

Best Local Similarity 97.1%; Pred. No. 0;
Matches 1181; Conservative 7; Mismatches 9; Indels 19; Gaps 2;

Oy 259 EAWNSLASSRSEQESTVWVOPPFETTPSILSEDPRQLQSTTLAWHMRMMPAGE 318

Db 1 EAWNSLASSRSEQESTVWVOPPFETTPSILSEDPRQLQSTTLAWHMRMMPAGE 60

Oy 319 KEDPLSYKHGRDKMKCPQSOPSPYLSYRNNSYJTRLOKPDQKU----- EVRE 364

Db 61 KEDPLSYKHGRDKMKCPQSOPSPYLSYRNNSYJTRLOKPDQKU----- :||| 120

Oy 365 GABIRCCKDKEPDTPVTPSVHRKPADINVIGALGSDTAGNGAGSTCGNVLIDVLTQVRL 424

Db 121 GABIRCCKDKEPDTPVTPSVHRKPADINVIGALGSDTAGNGAGSTCGNVLIDVLTQVRL 180

Oy 425 SWSVGGDENIGVTTNLANIRFENPSIJKGSFTSGIGKETSPNWFNQAVAGGRADLEVQA 484

Db 181 SWSVGGDENIGVTTNLANIRFENPSIJKGSFTSGIGKETSPNWFNQAVAGGRADLEVQA 240

Oy 485 RFLVDLMKNDTTHFQEDWKITLTFIGENDLDFCNLDLVHYRQPNTDNIGALDIHAE 544

Db 241 RRLVDLMKNDTTHFQEDWKITLFIGENDLDFCNLDLVHYRQPNTDNIGALDIHAE 300

Oy 545 VRAFVNLVTVBIVNRELRYKKVVKCPRMILSLCPVLFEDDISTELATLIEFKKF 604

Db 301 VRRAFVNLVTVBIVNRELRYKKVVKCPRMILSLCPVLFEDDISTELATLIEFKKF 360

Oy 605 QEKTHOLIESGYKDTRDFTVWQPFENVMDPKTSQGLPDSFAADCFFHSKSHSA 664

Db 361 QKTHOLIESGYKDTRDFTVWQPFENVMDPKTSQGLPDSFAADCFFHSKSHSA 420

Oy 665 ASALWNMMLEPVGQKTRHKFENKINTCPCNQVQPFRLTYK---- NSMQGIGTWLPCRD 719

Db 421 ASALWNMMLEPVGQKTRHKFENKINTCPCNQFWSLSTRLFDLYALRGHTWLPCRD 480

Oy 720 RAPSALHPTSVHARPADIOVVALGSLTANGIGKPPDDPDTYQGHSYSAGGD 779

Db 481 RAPSALHPTSVHARPADIOVVALGSLTANGIGKPPDDPDTYQGHSYSAGGD 540

Oy 780 SLBNVTUPLNIREFNLNTGIGVAGTGDDANDTAFNQAVPRAKAELIMSOVTLQMKK 839

Db 541 SLBNVTUPLNIREFNLNTGIGVAGTGDDANDTAFNQAVPRAKAELIMSOVTLQMKK 600

Oy 840 DDHRVNTHEDKVITVLLIGGSPLCDYCTDSNLYSAANFDURNALDVLRREVPRVJNL 899

Db 601 DDHRVNTHEDKVITVLLIGGSPLCDYCTDSNLYSAANFDURNALDVLRREVPRVJNL 660

Oy 900 VDFLNPTIMROVFLGNRDLCDYCTDSNLYSAANFDURNALDVLRREVPRVJNL 959

Db 661 VFLNPTIMROVFLGNRDLCDYCTDSNLYSAANFDURNALDVLRREVPRVJNL 720

Oy 960 GSGRYDQEDESVLQPFONQLPVADGDPSTSFPADCPCHPNQFRHSCALARLMNM 1019

Db 721 GSGRYDQEDESVLQPFONQLPVADGDPSTSFPADCPCHPNQFRHSCALARLMNM 780

Oy 1020 LEPLGSKETLDRAEMPTCONEPFPLPANSNTYPIKPAIENWGSDLTCTENKAS 1079

Db 788 LEPGSKETLDRAEMPTCONEPFPLPANSNTYPIKPAIENWGSDLTCTENKAS 840

Oy 1080 NSVPTSYHQLRADIKYVAAGDSLTTAVGARPARNNSSDLPTSWRGLWSIGGDNLEHT 1139

Db 841 NSVPTSYHQLRADIKYVAAGDSLTTAVGARPARNNSSDLPTSWRGLWSIGGDNLEHT 900

Oy 1140 TLPNLKCFNPYLIGFSTSTMEGTAGINVAEGARADMPACDWLTERMKNSPDINLEK 1199

Db 901 TLPNLKCFNPYLIGFSTSTMEGTAGINVAEGARADMPACDWLTERMKNSPDINLEK 960

Oy 1200 DWKLVLTIFGVNDLICHYCEPAAHLATEYVHQIQLAIDLSELPLPRAFVVMEASL 1259

Db 961 DWKLVLTIFGVNDLICHYCEPAAHLATEYVHQIQLAIDLSELPLPRAFVVMEASL 1020

Oy 1260 YOOGGKCKAMLAQNCTCLRHSSQSLLEKQELKVNMLQHGTSFSYWHQYTOREDPAV 1319

Db 1021 YOOGGKCKAMLAQNCTCLRHSSQSLLEKQELKVNMLQHGTSFSYWHQYTOREDPAV 1080

Oy 1320 VYOPPFONTLTPLNERSGDTDPLFSSBEDCPHFSDRGHAEMATLWNMLNLEPVKRTTSNF 1379

Db 1081 VYOPPFONTLTPLNERSGDTDPLFSSBEDCPHFSDRGHAEMATLWNMLNLEPVKRTTSNF 1140

Oy 1380 TISRAKUCKCPSPSPYLTNSRSLPQDAEPAEVLYWAVPVAAGGLVWQIGITWWR 1439

Db 1141 TISRAKUCKCPSPSPYLTNSRSLPQDAEPAEVLYWAVPVAAGGLVWQIGITWWR 1200

Oy 1440 CRRGGRRDPWMSLT 1455

Db 1201 CRRGGRRDPWMSLT 1216

RESULT 7
ID AAW30751
ID AAW30751 standard; Protein; 1450 AA.
XX
AC AAW30751;
XX
DT 28-JAN-1998 (first entry)
XX
DE Rat phospholipase-B/lipase.
XX
KW phospholipase B; lipase; supplement; pancreatic phospholipase; reagent;
XX
OS Rattus rattus.
XX
FT Key Location/Qualifiers
FT peptide 1..30
FT /label= signal_peptide
FT Protein 31
FT /label= mature_protein
FT Region 43..652
FT /label= repeat_region_1

| | | | | |
|----------|---|--|----|---|
| FT | Region | 367..712 | Db | 483 GLAAQAKKLVLASKDKTKTINFOBDWKITTVFGLNDLGCSCNLRSPQTFTDNKTL 542 |
| FT | /label= repeat_region_2 | 714..1059 | Oy | 539 DLHAEVPRAFNLVLTUENLRELYQEKWYCPMLRSLCPVLFKDINSTEALTI 598 |
| FT | Region | /label= repeat_region_3 | Db | 543 DLHAEVPRAFNLVLTUENLRELYQEKWYCPMLRSLCPVLFKDINSTEALTI 602 |
| FT | /label= repeat_region_4 | 1070..1408 | Oy | 599 EFNKKFOEKTQOLIESGRYDTRDFTVVQPFENDMVKTSSEGUDPNSFPAPDCHFSS 658 |
| FT | /label= transmembrane_binding_domain | 1421..1443 | Db | 603 ENRQVQETGKLGIESGRYDTRDFTVLQLPTEFLPFRVSCPMLRLCPVLTNLGNSAELAQLV 662 |
| FT | Domain | XX | Oy | 663 KTHARSAALWKNMLPEVGOKTTRHKPEENKNTTCRNQVQPLRTYKNSMCGHTWLPCR 718 |
| PD | 22-SEB-1997. | XX | Db | 663 KTHARSAALWKNMLPEVGOKTTRHKPEENKNTTCRNQVQPLRTYKNSMCGHTWLPCR 722 |
| PF | 15-MAR-1996; | 96JP-0086022. | Oy | 719 DRAPSALHPTSHALKRADIQVWAAGDSLUTAGNGGSKPPDLPLDTTQXGLSYAGGD 778 |
| PR | 15-MAR-1996; | 96JP-0086022. | Db | 723 EKAPSASPTSTHALKRADIQVWAAGDSLUTAGNGGSKPPDLPLDTTQXGLSYAGGD 782 |
| PA | (TOJO/) TOKIO H. | PA (TOFO/) TOKIO H. | Oy | 779 GSLENVTILPNTLREFNRNLTGYAVGTDANDTNAFAQANQAVPGAKADELMSQVQTLMQRM 838 |
| PA | XX | PS | Db | 843 KNDTRVNPHQDMRKVITVMIGASDLCDCDKDSNRYSANFSDHRLNALDILKEVPRALVN 902 |
| DR | N-PSDB; AAT91874. | XX | Oy | 899 LVDPLNITIMRQVFLGPDKPCPVQQSVALCVCYLURENSGELARLEAFSHAYRSEBEL 958 |
| PT | DNA encoding protein with phospholipase B and lipase activity - useful as supplement to pancreatic phospholipase | PS | Db | 903 LYDFMNNTSIIRQVPLKDFKPCPVQNTSVLCLCVLTCQFGBSDHLARLEAFTHYSQESNLQ 962 |
| PT | useful as supplement to pancreatic phospholipase | CC | Oy | 959 VSGRYVQEDTSVLIQPFONIQLPVLAQDGLPDTSPAPDCIHPNQKHSOLARALWTN 1018 |
| XX | Claim 1; Pages 7-12; 16pp; Japanese. | CC | Db | 963 VSGRYVTREDSVLQFPLRIPLENQMPDTSFFAPOILLSQFHQLARALWN 1022 |
| XX | The present sequence is a protein which has both phospholipase B and pancreatic phospholipase and as a reagent for the determination of phospholipase B/lipase. | CC | Oy | 1019 MIRPLPGKTTETIDIRASPMTCPTOHEPLRMPTRNSNTYPIKPAIENWGSDFLCTEWKA 1078 |
| Sequence | 1450 AA; | CC | Db | 1023 MBLPLGKMDTLDPKELIALACPTKOPFLRPTFRNSNTYPIKPAIENWGSDFLCTEQSP 1082 |
| Query | Match | 68.0%; Score 5279.5; DB 18; Length 1450; | Oy | 1079 SNSVPTSHQLRADITKVAALGDSLTAVERARPNNSSDLPSPWRLSWSFGGDGMLETH 1138 |
| Query | Best Local Similarity | 70.0%; Pred. No. 0; | Db | 1083 SSKVPTSHQLRADITKVAAMDPSSLDPTPWRLSWSFGGDGMLETH 1142 |
| Matches | 1012; | Mismatches 161; Indels 9; Gaps 6; | Oy | 1139 TLPNIKTFKNYLGLGSTSWEGTAGLNVAEGARARDMPAQAWDIVERMKSPDNLNE 1198 |
| Oy | 5 PGJIFL..-LELLIAQGQTPQINTSPRSKSTLE3GQIWPETLKNQFPFCPNKLGYNMRSV | 62 | Db | 1143 TLPNIKTFKNYLGLGSTSWEGTAGLNVAEGARARDMPAQAWDIVERMKSPDNLNE 1202 |
| Db | 5 PGJIFL..-LELLIAQGQTPQINTSPRSKSTLE3GQIWPETLKNQFPFCPNKLGYNMRSV | 62 | Oy | 1199 KOMKLVITFIGNDLCHYCENCENEAHLATEYQHQIQLADITISEELPRAFVNVEMLAS 1258 |
| Oy | 63 HS1KPSDKTFKVAIGNLTPQVQFVQVPLKPSVQVMEKPSOSLSELQNLVQIGMTALSDIRHNP | 124 | Db | 1203 EDWKLTITLGNNDLCLCVDENPYSTRERYKVIHALDIFVEELPRAFVNVEMLAS 1262 |
| Db | 63 HS1KPSDKTFKVAIGNLTPQVQFVQVPLKPSVQVMEKPSOSLSELQNLVQIGMTALSDIRHNP | 124 | Oy | 1259 LYQOGCKACM-LAQNNTCTCRHSOSSEKELKQWNNLQHGIISSFSYHMQYTORBD 1317 |
| Oy | 120 SVMPVQHGTGKRVIPHQD-AEIDWIKQOELVRNKMENQDQFDWKLNINVUPNSASOCY | 178 | Db | 1263 LWDQGGCICAMPALAVQKNCISKRSQIMAMOBKKYNGNQSOALSBLSYHMRWEDP 1322 |
| Db | 125 SVIMPCTSPKGKSTAGHTTIAELWIKQAKELVRLHDKNPFEDRDKWLTIVFSNTSQC | 184 | Oy | 1318 AVVQQPFPNTLPLNIEGDDTLPFSEDCFHSFSDRGHAEMIALMMNMLPVGKRTSN 1377 |
| Oy | 179 LCTCSAQQNLAGAGVDELMGVLDYQLOQEVPRATVNLVLDSEAEVSROYHGWLSPAGEP | 238 | Db | 1323 AVVQQPFPNTLPLNIEGDDTLPFSEDCFHSFSDRGHAEMIALMMNMLPVGKRTSN 1382 |
| Db | 185 LCSSDQQRRHMHGMEMISMGVLUYLHRVPRAVNLVLDSEULMAQHQEOTSPAPEI | 244 | Oy | 1378 NFTSRMALKCSCPESPYLTYLNRSLRIDPQABEAPVLYWVPAAGVGLVWGLICVV 1437 |
| Oy | 239 CNGSEBETRILAKYUNQSYOEAANSLASSRSSEOESETVRQFPRFVTPSHSEPRL | 298 | Db | 1383 NFTYNRTRKLCSCPENPYFLYTVRNSQQLDRAKAKENSNTLYAWPVAAGGLVGLGML 1442 |
| Db | 245 CKSEBEITKLSKAVMQWSYQEAEDLASSKFNKHETPFAVVFOSFFSEV-ELPLERSP | 302 | Oy | 1418 WRCRR 1442 |
| Oy | 299 QDSTLAWHLNTRMMPAGEKDPPLSVHGRPMKCPSESPPVFSYRNSVYTRQRPD | 358 | Db | 1443 WRTRW 1447 |
| Db | 303 QDSTLALRINSMNPEPGKRGKDTLNEARKTMKCPSESPVLYPTYNSYCAROLKIG | 362 | Oy | RESULT 8 |
| Oy | 359 KLEVREGABAIRCPDXPSDTPVTPVHRAKPDADIVGALGDSITAGNAGSTPGNVLDL | 418 | Db | AAB34448 |
| Db | 363 KFQMKEGKFKTCPDKDPSIPTVHRAKPDADIVGALGDSITAGNAGSTPGNVLDL | 422 | Oy | AAB34448 standard; Protein: 969 AA. |
| Oy | 419 TORGLSLMSVGPGNIGTWTNLREFNPSLKGFSVGTGKETSPNPLNORVAGRAE | 478 | Db | AAB34448 |
| Db | 423 TORGLSLMSVGQDGTETVTLANLREFNPSLKGFSVGTGKETSPNQAVAGAKSD | 482 | Oy | 479 DLPUQARRLVLDMQNDTRIHFPQDMKIIILFIGNDICDFCNLVLVHSPQNFNDNIGKAL 538 |

| | |
|--|---|
| XX | PT polymucleotides, useful for diagnosing, treating or preventing |
| DT 14-MAY-2003 (first entry) | PT cardiovascular disorders (e.g. aneurysm), neurological disorders (e.g. |
| XX | PT Parkinson's disease) or cancers - |
| DE | |
| XX | |
| KW Human; lipid-associated molecule; LIPAM; cardiovascular disorder; stroke; | |
| KW arteriovenous fistula; atherosclerosis; hypertension; Raynaud's disease; | |
| KW aneurysm; congeitive heart failure; thromboophlebitis; angina pectoris; | |
| KW ischaemic heart disease; rheumatic heart disease; peptic oesophitis; | |
| KW gastrointestinal disorder; lipid metabolism disorder; Crohn's disease; | |
| KW nausea; peptic ulcer; fatty liver; Fabry's disease; Gaucher's disease; | |
| KW diabetes mellitus; hyperlipidemia; hypercholesterolaemia; epilepsy; | |
| KW auto-immune disorder; inflammatory disorder; neurological disorder; kuru; | |
| KW acquired immunodeficiency syndrome; anaemia; Alzheimer's disease; asthma; | |
| KW dementia; prion disease; Creutzfeldt-Jakob disease; leukaemia; cancer; | |
| KW adenocarcinoma; lymphoma; melanoma; myeloma; Barcoma; gene therapy; | |
| KW protein replacement therapy. | |
| OS Homo sapiens. | |
| XX | |
| KEY Location/qualifiers | |
| FH Peptide 1..19 | |
| FT /label= Signal-peptide | |
| FT Protein 20..969 | |
| FT /note= "Human mature LIPAM-9 protein" | |
| FT Peptide 1..21 | |
| FT /label= Signal-peptide | |
| FT Protein 22..969 | |
| FT /note= "Human mature LIPAM-9 protein" | |
| FT Peptide 1..22 | |
| FT /label= Signal-peptide | |
| FT Protein 23..969 | |
| FT /note= "Human mature LIPAM-9 protein" | |
| FT Peptide 1..25 | |
| FT /label= Signal-peptide | |
| FT Protein 26..969 | |
| FT /note= "Human mature LIPAM-9 protein" | |
| FT Domain 393..521 | |
| FT /note= "Lipase/acetylhydrolase with GDSL-like motif" | |
| FT Active-site 394..404 | |
| FT /note= "Lipolytic enzyme G-D-S-L Family | |
| FT Serine active-site 740..868 | |
| FT /note= "Lipase/acetylhydrolase with GDSL-like motif" | |
| FT Active-site 741..751 | |
| FT /note= "Lipolytic enzyme G-D-S-L family | |
| FT Serine active-site" | |
| PN WO200394983-A2. | |
| XX | |
| PD 28-NOV-2002. | |
| XX | |
| PF 17-MAY-2002; 2002WO-US15688. | |
| XX | |
| PR 18-MAY-2001; 2001US-292242P. | |
| PR 25-MAY-2001; 2001US-293726P. | |
| PR 01-JUN-2001; 2001US-295346P. | |
| PR 06-JUL-2001; 2001US-303404P. | |
| PR 24-AUG-2001; 2001US-314754P. | |
| PR 22-JAN-2002; 2002US-351262P. | |
| PR 29-MAR-2002; 2002US-368799P. | |
| XX | |
| PA (INCYT- INCYTE GENOMICS INC. | |
| XX | |
| PI Tang YT, Yue H, Azimzai Y, Baughn MR, Burford N, Reddy R, Gandhi AR; | |
| PI Wallia NK, Das D, Nguyen DB, Yao MG, Arvizu CS, Lu Y, Griffin JA, Elliott VS, Ramkumar J, Lal PG, Lu DAM, Lee EA; | |
| PI Lee SY, Yue H, Yang J, Tribouley CM, Kable AE, Swarnakar A; | |
| DR N-PSDB; AAD52634. | |
| PT New human lipid-associated molecule (LIPAM) proteins and | |
| PT polymucleotides, useful for diagnosing, treating or preventing | |
| PT cardiovascular disorders (e.g. aneurysm), neurological disorders (e.g. | |
| PT Parkinson's disease) or cancers - | |
| XX | |
| PS Claim 64; Page 153-156; 171pp; English. | |
| XX | |
| CC The present invention relates to novel human lipid-associated molecules | |
| CC (LIPAM) and polymucleotides encoding such proteins. Sequences of the | |
| CC invention are useful for treating diseases or conditions associated with | |
| CC decreased expression of functional LIPAM. The antagonist is useful for | |
| CC treating a disease or condition associated with the overexpression of | |
| CC functional LIPAM. They are useful for diagnosing, treating or preventing | |
| CC cardiovascular disorders (e.g. arteriovenous fistula, atherosclerosis, | |
| CC hypertension, Raynaud's disease, aneurysm, varicose veins, congestive | |
| CC heart failure, thromboophlebitis, angina pectoris, ischaemic heart disease | |
| CC or rheumatic heart disease), gastrointestinal disorders (e.g. peptic ulcer, | |
| CC oesophagitis, nausea, peptic ulcer or Crohn's disease), lipid metabolism | |
| CC disorders (e.g. hyperlipidemia, hypercholesterolaemia), autoimmune disorders | |
| CC or inflammatory disorders (e.g. acquired immunodeficiency syndrome, | |
| CC epilepsy, dementia, Alzheimer's disease, or prion diseases such as kuru, | |
| CC or Creutzfeldt-Jakob disease) or cancers (e.g. adenocarcinoma, leukaemia, | |
| CC lymphoma, melanoma, myeloma or sarcoma). They are also used in gene | |
| CC therapy and protein replacement therapy. The present sequence is human | |
| XX | |
| SQ Sequence 969 AA; | |
| Query Marcch Best Local Similarity 63.2%; Score 4911; DB 24; Length 969; | |
| Matches 923; Conservative 0; Mismatches 1; Indels 0; Gap 0; | |
| QV 1 MGLRPGLPFLUELLLGCGTPOINTSPRKSTKLEQGQMPETLKNSPFPFCNPNIKLGVNMPSK 60 | |
| Db 1 MGLRPGLPFLUELLLGCGTPOINTSPRKSTKLEQGQMPETLKNSPFPFCNPNIKLGVNMPSK 60 | |
| QV 61 SVHSLSKPSDKIIVKAIGNLEIIPDPGICDLEQDWTRPQQVCMGMVNTVLSDTIRYSPS 120 | |
| Db 61 SVHSLSKPSDKIIVKAIGNLEIIPDPGICDLEQDWTRPQQVCMGMVNTVLSDTIRYSPS 120 | |
| QV 121 VPMVPCITGKRVIPHDGAEDWIAQAEVLVRNNKENIQDFDPMWKLNVNFSNASOYL 180 | |
| Db 121 VPMVPCITGKRVIPHDGAEDWIAQAEVLVRNNKENIQDFDPMWKLNVNFSNASOYL 180 | |
| QV 181 PSAQONGLAGSYDLEMVLDLQOEPPRFALNVLDSVEAVRSRQHTNLSAPBCCN 240 | |
| Db 181 PSAQONGLAGSYDLEMVLDLQOEPPRFALNVLDSVEAVRSRQHTNLSAPBCCN 240 | |
| QV 241 CSEBETTRAKVUNMOWSYQABNNSLLASSRYSBQESPTVVFQFFYETTPSLJASEDPRLD 300 | |
| Db 241 CSEBETTRAKVUNMOWSYQABNNSLLASSRYSBQESPTVVFQFFYETTPSLJASEDPRLD 300 | |
| QV 301 STCLAWMLWLNRMWEPAGEKDEBLSVKRGPRMKCPSDPSYPSYPSYRISNLTRQPKDKL 360 | |
| Db 301 STCLAWMLWLNRMWEPAGEKDEBLSVKRGPRMKCPSDPSYPSYRISNLTRQPKDKL 360 | |
| QV 361 EYREGAERCPDKDPSDPTVPSVHLKPADINVIGALGDSLTAGNGAGSTGNVLDLQ 420 | |
| Db 361 EYREGAERCPDKDPSDPTVPSVHLKPADINVIGALGDSLTAGNGAGSTGNVLDLQ 420 | |
| QV 421 YRGLSWSGGDENIGTVTLANTIREFPLSKGFSVGTGKTSPTNAFLQAVAGRABDL 480 | |
| Db 421 YRGLSWSGGDENIGTVTLANTIREFPLSKGFSVGTGKTSPTNAFLQAVAGRABDL 480 | |
| QV 481 PYQARRVMDMONTRHFOBDWKIITFLIGENDLCFCNDLVHYPQRNFUDNIGRADI 540 | |
| Db 481 PYQARRVMDMONTRHFOBDWKIITFLIGENDLCFCNDLVHYPQRNFUDNIGRADI 540 | |
| QV 541 LHAEVPRAFVNUVLUVNLRELYQKVKYCPMLSLCPVLUKFDPNSTELATLIEF 600 | |
| Db 541 LHAEVPRAFVNUVLUVNLRELYQKVKYCPMLSLCPVLUKFDPNSTELATLIEF 600 | |

QY 601 NKKFQEKITHQLIISCGRYTREDFWVQPPFENDMPKTSCEGLPDNSRAPCPHRSSKS 660
 CC CC models or targets for the development of human therapeutic agents that
 Db 601 NKKFQEKITHQLIISCGRYTREDFWVQPPFENDMPKTSCEGLPDNSRAPCPHRSSKS 660
 CC module phospholipase activity in cells and tissues that express the
 CC phospholipase, such as in kidney, blood, lung, brain glioblastomas,
 CC prostate, colon or leukocytes.
 XX SQ Sequence 472 AA:
 Query Match 32.1%; Score 2492; DB 23; Length 472;
 Best Local Similarity 93.7%; Pred. No. 2e-216; Matches 472; Conservative 0; Mismatches 0; Indels 32; Gaps 1;
 Matches 472; Conservative 0; Mismatches 0; Indels 32; Gaps 1;
 Db 721 APSALHPTSVHLRPAIDIQVALGSLSLTAGNGIGSKPFDLDPVTQYRGYSAGGDS 780
 QY 781 LENTTILNIRLEPNRNTGAYCTGDAINTFLNQAVPGACEDMSQVQTIMQKRD 840
 Db 781 LENTTILNIRLEPNRNTGAYCTGDAINTFLNQAVPGACEDMSQVQTIMQKRD 840
 QY 841 DHRVNPHEDWKVITVLIGSDCDYCTSNSLYSAANFVHRLNALDVLAVEREVLVNLV 900
 Db 841 DHRVNPHEDWKVITVLIGSDCDYCTSNSLYSAANFVHRLNALDVLAVEREVLVNLV 900
 QY 901 DFLNPTIMQVFLGPDKCPVQOA 924
 Db 901 DFLNPTIMQVFLGPDKCPVQOA 924

RESULT 9

ID ABP53556 standard; Protein; 472 AA.

XX AC ABP53556;

XX DR 16-DEC-2002 (first entry)

XX DE Human phospholipase protein SEQ ID NO:2.

XX KW Human; phospholipase; enzyme; chromosome 2.

XX OS Homo sapiens.

XX PN WO200202977-A2.

XX PD 15-AUG-2002.

XX PF 28-JAN-2002; 2002WO-US02302.

XX PR 08-FEB-2001; 2001US-0778961.

XX PA (PEKE) PE CORP NY.

XX PT Yan C, Ketchum KA, Di Francesco V, Beasley EM;

XX DR WPI; 2002-682698/73.

DR N-PSDB; ABQ82234.

XX PT New human phospholipase proteins useful for the development of human
 PT therapeutics and diagnostic compositions, drug screening assays, tissue
 PT typing and pharmacogenomic analysis -

XX PS Claim 1; Fig 2A; 95pp; English.

CC The present sequence represents a human phospholipase protein (1) located
 CC on chromosome 2. (1) can be used for identifying agents that modulate its
 CC function or activity where the agent is useful for treating a disease
 CC or condition mediated by a the human phospholipase protein. (1) peptides
 CC can be used in substantial and specific assays related to functional
 CC information of the peptide sequences, to raise antibodies or to elicit
 CC immune response, as reagents in assays that determine the levels of
 CC protein in biological fluids, and as markers for tissues where the
 CC corresponding protein is expressed. Nucleotide sequences encoding (1)
 CC can be used as probes, primers and chemical intermediates in biological
 CC assays, for constructing recombinant vectors, and expressing antigenic
 CC portions of the protein. (1) and nucleic acid molecules encoding it can
 CC be used in the identification of therapeutic proteins and may serve as

QY 955 MRELVGSRDQTQEDFSVTLQPFONLQLPVADGLPDTSFAPDCIHNPQKPHSQJARA 1014
 Db 1 MRELVGSRDQTQEDFSVTLQPFONLQLPVADGLPDTSFAPDCIHNPQKPHSQJARA 1014
 QY 1015 LWTMNLBEGSKETULRAEMPTCTQNEPLRTNRNTYPIKAENWGSDECT 1074
 Db 34 -----LELGSKETULRAEMPTCTQNEPLRTNRNTYPIKAENWGSDECT 88
 QY 1075 EWKAQNSNPTSHQLPADIKVVAALGDSLTTAVGARNNSDLPTSWRGLSWSIGGDN 1134
 Db 89 EWKAQNSNPTSHQLPADIKVVAALGDSLTTAVGARNNSDLPTSWRGLSWSIGGDN 148
 QY 1115 LEHTTILNILKKNPFLIGPSTSTWEGTAGLNVAARGARADMPAQDWLVMQNSPD 1194
 Db 149 LEHTTILNILKKNPFLIGPSTSTWEGTAGLNVAARGARADMPAQDWLVMQNSPD 208
 QY 1115 INLEKDWLKLUTLPIGNNDLCHYCNPEAHATEYQHINQADILSSELPRAFVNUEM 1254
 Db 209 INLEKDWLKLUTLPIGNNDLCHYCNPEAHATEYQHINQADILSSELPRAFVNUEM 268
 QY 1255 ELASLYCQGGKCAMLAQNCTCLRHQSSEKQELKVNVLORGTSFSYWHQYQR 1314
 Db 259 ELASLYCQGGKCAMLAQNCTCLRHQSSEKQELKVNVLORGTSFSYWHQYQR 328
 QY 1315 EDFAVVVQPFQONTLTPNERGPTDIAFPSEDCPHFSRDRGAEMATALLNNMLEPVGRT 1374
 Db 329 EDFAVVVQPFQONTLTPNERGPTDIAFPSEDCPHFSRDRGAEMATALLNNMLEPVGRT 388
 QY 1375 TSNMPHTSRAKLKICPSPSPPYIYLTRSLIDQAEAPEAEPYVAVWVAGYGLWING 1434
 Db 339 TSNMPHTSRAKLKICPSPSPPYIYLTRSLIDQAEAPEAEPYVAVWVAGYGLWING 448
 QY 1435 TVWRCRGGRRDPPMSLRTVAL 1458
 Db 449 TVWRCRGGRRDPPMSLRTVAL 472

RESULT 10

ID ABU65082 standard; Protein; 310 AA.

XX AC ABU65082;

XX DR 20-MAY-2003 (first entry)

XX DE Human Nov24b protein.

XX PT NOVK; cytostatic; cardiotropic; antiarteriosclerotic; antiasthmatic; cancer;
 KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KW human.

XX OS Homo sapiens.

XX PN WO200212757-A2.

XX PD 19-SEP-2002.

XX PR 08-MAR-2002; 2002WO-US06908.

XX PR 08-MAR-2001; 2001US-274101P.

PR 08-MAR-2001; 2001US-274194P.

PR 08-MAR-2001; 2001US-274281P.

| | | | |
|----|--|-----------------------|--|
| PR | 08-MAR-2001; 2001US-274322P. | PT | treating a disorder associated with aberrant NOVX expression or |
| PR | 09-MAR-2001; 2001US-274849P. | PT | activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or |
| PR | 12-MAR-2001; 2001US-275235P. | PT | bronchial asthma - |
| PR | 13-MAR-2001; 2001US-275578P. | XK | |
| PR | 13-MAR-2001; 2001US-275579P. | PS | |
| PR | 13-MAR-2001; 2001US-275601P. | XK | |
| PR | 14-MAR-2001; 2001US-27600P. | CC | Claim 1; Page 169; 1103pp; English. |
| PR | 16-MAR-2001; 2001US-276776P. | CC | This invention describes novel human NOVX polypeptides which have |
| PR | 19-MAR-2001; 2001US-276994P. | CC | cytostatic; cardiant; anti-arteriosclerotic; antiasthmatic and |
| PR | 20-MAR-2001; 2001US-277239P. | CC | hypotensive activity. Pharmaceutical compositions comprising the NOVX |
| PR | 20-MAR-2001; 2001US-277321P. | CC | proteins or nucleic acid molecules or NOVX antibodies are useful for |
| PR | 20-MAR-2001; 2001US-277327P. | CC | preventing or treating a disorder associated with aberrant NOVX |
| PR | 21-MAR-2001; 2001US-27791P. | CC | expression or activity e.g.; cancer, hypertension, atherosclerosis, |
| PR | 22-MAR-2001; 2001US-277833P. | CC | cardiomyopathy or bronchial asthma. The products of the invention can |
| PR | 23-MAR-2001; 2001US-278152P. | CC | be used for gene therapy or in a vaccine. ABU65041-ABU65218 represent |
| PR | 26-MAR-2001; 2001US-278999P. | CC | the NOVX polypeptides encoded by ABX97008-ABX97185. |
| PR | 27-MAR-2001; 2001US-27936P. | XK | |
| PR | 28-MAR-2001; 2001US-27934P. | SQ | |
| PR | 30-MAR-2001; 2001US-277338P. | Sequence | 310 AA; |
| PR | 30-MAR-2001; 2001US-27935P. | Query | Match |
| PR | 30-MAR-2001; 2001US-280233P. | Best Local Similarity | 19.6%; |
| PR | 02-APR-2001; 2001US-280802P. | Matches | Score 1525; |
| PR | 02-APR-2001; 2001US-280900P. | Loc | DB 23; |
| PR | 04-APR-2001; 2001US-281194P. | Db | Length 310; |
| PR | 13-APR-2001; 2001US-283675P. | Qy | |
| PR | 30-APR-2001; 2001US-287424P. | 491 | MQNDRTRHFQEQWKMITLFIGNDLCPDNQDLYHSQNPNTDMIGKALDIHLAEPRAVY |
| PR | 02-MAY-2001; 2001US-28866P. | Db | 550 |
| PR | 03-MAY-2001; 2001US-28834P. | Qy | NUTVLTENLNLRLYQESKKVCPRMILRSCLCPVLFKDDNISTELATLIEFNKKFQEKTQ |
| PR | 03-MAY-2001; 2001US-288528P. | Db | 610 |
| PR | 15-MAY-2001; 2001US-291190P. | Qy | LIESGRDTREDFTVQVQPFPPENNDMKTSEGPDNSFPAPCOPFHESKSRSBASALNN |
| PR | 16-MAY-2001; 2001US-291199P. | Db | 670 |
| PR | 16-MAY-2001; 2001US-291240P. | Qy | LIESGRDTREDFTVQVQPFPPENNDMKTSEGPDNSFPAPCOPFHESKSRSBASALNN |
| PR | 30-MAY-2001; 2001US-294485P. | Db | 180 |
| PR | 31-MAY-2001; 2001US-294889P. | Qy | NMLEPVQGKTRHKPEKNINTCPNQVQPRLTYKNSMQGHGWLPCDRAPSALHPSV |
| PR | 31-MAY-2001; 2001US-294999P. | Db | 730 |
| PR | 18-JUN-2001; 2001US-299027P. | Qy | HALRPADIOVVAALGDSLTTAGNGIGSKRDPDPVTTOYRLGSYSGGDGSLEN |
| PR | 19-JUN-2001; 2001US-299310P. | Db | 731 |
| PR | 10-JUL-2001; 2001US-304354P. | Qy | HALRPADIOVVAALGDSLTTAGNGIGSKRDPDPVTTOYRLGSYSGGDGSLEN |
| PR | 31-JUL-2001; 2001US-309198P. | Db | 730 |
| PR | 16-AUG-2001; 2001US-312903P. | Qy | HALRPADIOVVAALGDSLTTAGNGIGSKRDPDPVTTOYRLGSYSGGDGSLEN |
| PR | 10-SEP-2001; 2001US-318462P. | Db | 731 |
| PR | 12-SEP-2001; 2001US-318770P. | Qy | HALRPADIOVVAALGDSLTTAGNGIGSKRDPDPVTTOYRLGSYSGGDGSLEN |
| PR | 27-SEP-2001; 2001US-325430P. | Db | 730 |
| PR | 27-SEP-2001; 2001US-325681P. | Qy | HALRPADIOVVAALGDSLTTAGNGIGSKRDPDPVTTOYRLGSYSGGDGSLEN |
| PR | 18-OCT-2001; 2001US-330380P. | Db | 731 |
| PR | 31-OCT-2001; 2001US-335301P. | Qy | HALRPADIOVVAALGDSLTTAGNGIGSKRDPDPVTTOYRLGSYSGGDGSLEN |
| PR | 14-NOV-2001; 2001US-332172P. | Db | 730 |
| PR | 14-NOV-2001; 2001US-332271P. | Qy | HALRPADIOVVAALGDSLTTAGNGIGSKRDPDPVTTOYRLGSYSGGDGSLEN |
| PR | 14-NOV-2001; 2001US-333184P. | Db | 731 |
| PR | 14-NOV-2001; 2001US-333272P. | Qy | HALRPADIOVVAALGDSLTTAGNGIGSKRDPDPVTTOYRLGSYSGGDGSLEN |
| PR | 21-NOV-2001; 2001US-332094P. | Db | 730 |
| PR | 03-DEC-2001; 2001US-337426P. | Qy | HALRPADIOVVAALGDSLTTAGNGIGSKRDPDPVTTOYRLGSYSGGDGSLEN |
| PR | 03-DEC-2001; 2001US-338092P. | Db | 731 |
| PR | 04-DEC-2001; 2001US-337185P. | Qy | HALRPADIOVVAALGDSLTTAGNGIGSKRDPDPVTTOYRLGSYSGGDGSLEN |
| PR | 03-JAN-2002; 2002US-345705P. | Db | 730 |
| PR | 07-MAR-2002; 2002US-0092900. | Qy | HALRPADIOVVAALGDSLTTAGNGIGSKRDPDPVTTOYRLGSYSGGDGSLEN |
| PA | (CURA-) CURAGEN CORP. | XX | 293 |
| XX | | XX | |
| PI | Padiyar M, Spyrek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L; | XX | RESULT 11 |
| PI | Zerhouni BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R; | XX | ABB11053 |
| PI | Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernov V; | XX | ID ABB11053 standard; peptide; 267 AA. |
| PI | Fernandes ER, Casman SJ, Malyankar IM, Gerlach V, Liu Y; | XX | XX |
| PI | Anderson D, Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H; | XX | AC ABB11053; |
| DR | Alsobrook JP, Lepley DM, Rieger DK; | XX | XX |
| DR | WPI; 2002-723332/78. | XX | DT 11-JAN-2002 (first entry) |
| DR | N-PSDB; ABX97049. | XX | DE Human phospholipase B homologue, SEQ ID NO:1423. |
| XX | | XX | KW Human; cytokine; cell proliferation; cell differentiation; growth factor; |
| XX | | XX | KW haemopoiesis regulation; tissue growth; immunomodulator; activin; |
| XX | | XX | KW inhibitor; chemotaxis; chemokines; thrombolytic; oncogene; |
| XX | | XX | KW proliferation; metastasis; cancer; tumor; haemopoietic disorder; |
| XX | | XX | KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis; |
| XX | | XX | KW chronic inflammatory condition; proliferative retinopathy; |
| XX | | XX | KW atherosclerosis; coronary heart disease; arterial ischaemia; |
| XX | | XX | KW bone disorder; osteoporosis; vascular growth disorder; |
| XX | | XX | KW tissue regeneration; wound healing; infection; immune disorder; |
| XX | | XX | KW cell culture; drug screening; gene therapy; anti-inflammatory; |
| XX | | XX | KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic; |
| XX | | XX | KW cytotoxic; osteopathic; vasotropic; cardiant; virucide; antibacterial; |
| XX | | XX | KW antifungal; vulnerary; antileuker. |
| OS | Homo sapiens. | OS | |
| XX | | XX | WO2001157108-A2. |
| XX | | XX | 09-AUG-2001 |

| | | | |
|-------------|---|---|----|
| XX | 45 | -----MLELGSKTEFLDRAEMPTCPTQNEPFLRTPRNNSVYPI | 86 |
| PF | 05-FEB-2001; 2001WO-US03800. | | |
| XX | 1061 KPAIENKSDFLCTEWKSNSTPSVQIQRADIKVVALGSPSLTAVGARNPNSDPLT | 1120 | |
| PR | 03-FEB-2000; 2000US-0496914. | | |
| PR | 27-APR-2000; 2000US-0560875. | | |
| PA | (HYSE-) HYSEQ INC. | | |
| XX | WPI; 2001-457740/49. | | |
| PT | Tang YT, Liu C, Drmanac RT; | | |
| XX | DR | | |
| XX | N-PSDB; ABA8297. | | |
| PT | Human proteins and DNA encoding sequences useful for preventing, treating or ameliorating a medical condition in a mammalian subject e.g. arthritis and cancer - | | |
| XX | | | |
| PS | Claim 20; Page 140; 1963pp; English. | | |
| XX | | | |
| CC | Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; hematopoiesis regulatory activity; tissue growth activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis, depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, bacterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cell damaged by illness, autoimmunity or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and drug screening techniques. The present sequence represents a novel human polypeptide of the invention. | | |
| XX | sequence 267 AA; | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| Db | 45 ----- | 44 | |
| Qy | 1001 IHRNOKFHSQALARALWNTMLPLPGSKTEFLDRAEMPTCPTONEPELRTPRNSNYTPI | 1060 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 940 | |
| Db | 1 LRVNLDVLUHREPVRLVNLVDFNPTIMQVFLGNPDCKPQVQASVLNCVLTRENQSE | 44 | |
| Qy | 941 LARLEAFSRAYRSMSMRELVGSGRYDTQDFSVVLPQQFFQNQICQLPVLADGLPDTSFAPDC | 1000 | |
| SQ | | | |
| Query Match | 17.5%; Score 1360; DB 22; Length 267; | | |
| Matches | 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1; | | |
| Qy | 881 LRVNLDVLUHREPVRLVNLVDF | | |

CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antiandamic; anti-laggregant; haemostatic; pulmonary;
 CC antirheumatic; osteopathic; dermatologic; anti-allergic; anti-asthmatic;
 CC anti-parkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense and vaccine
 production. The Proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 rheumatoïd arthritis, septic shock, pancreatitis, cardiac dysfunction,
 neuropathology, cardia anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorder, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.

XX Sequence 267 AA;

Query Match 17.5%; Score 1360; DB 22; Length 267;
 Best Local Similarity 73.7%; Pred. No. 3.7e-114; Mismatches 0; Indels 94; Gaps 1;
 Matches 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1;

Qy 881 LRVNALDVVLHREVRVRLVNLVDLNFNPTIMQVFLGNPDKCPVQOASVLCNCVLTRENSE 940
 Db 1 LRVNALDVVLHREVRVRLVNLVDLNFNPTIMQVFLGNPDKCPVQO----- 44

Qy 941 LARLEAFSRAYRSSMRELVGSGRYDQEDPSVVLQPFONICQVPLADGGLPDTSSFFADDC 1000
 Db 45 ----- 44

Qy 1001 IHPNQKPHSQRALARLWNTMLEPLGSKTEITLDRAREMPITCPTONEPELRTPRNSNYPI 1060
 Db 45 ----- MLEPLGSKTEITLDRAREMPITCPTONEPELRTPRNSNYPI 86

Qy 1051 KPAIENKSDFLCTEWKASNSVPTVSHQLRPAKIVVALGSLTTAVGARNNSDLP 1120
 Db 87 KPAIENKSDFLCTEWKASNSVPTVSHQLRPAKIVVALGSLTTAVGARNNSDLP 146

Qy 1121 SWIGLGSWSIGGGDSNLETHTLPLNTLKKENPYLUGFSTSWEGTAGLNVAEAGARADMPA 1180
 Db 147 SWIGLGSWSIGGGDSNLETHTLPLNTLKKENPYLUGFSTSWEGTAGLNVAEAGARADMPA 206

Qy 1181 QAWDLVERMKNSPDINLKDWKLVTLFIGNDLHYCENPAAALATEVQHQIQQALDILS 1240
 Db 207 QAWDLVERMKNSPDINLKDWKLVTLFIGNDLHYCENPAAALATEVQHQIQQALDILS 266

Qy 1241 E 1241
 Db 267 E 267

RESULT 13

ID ABB71556 standard; Protein; 424 AA.

XX ABB71556;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 41460.

XX ID ABB9611
 XX ABB9611 standard; Protein; 148 AA.

AC ABB9611;

XX OS Drosophila melanogaster.

DT 21-JUN-2002 (first entry)

DE Human testicular antigen SEQ ID NO: 1495.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX Sequence 424 AA;

Query Match 7.5%; Score 584; DB 22; Length 424;
 Best Local Similarity 38.5%; Pred. No. 2.3e-43; Mismatches 116; Indels 46; Gaps 12;
 Matches 137; Conservative 57; Mismatches 116; Indels 46; Gaps 12;

Qy 369 RPPDKDSDSTVTSVHLKPADINVAGLGSLTAGNGAGSPTPGNVLDVLTQYRGLSWSV 428
 Db 88 RSPER-----PTSVHLRPGDIDVIGAMGDSLTTAGNGIFAT--NLLHVTVENRGVMSI 139

Qy 429 GGDENIGTVTTLANLIRPFNFSLKGSGVGTKETSPNAFLNOAVAGGRADLPVQARLV 488
 Db 140 GQYDPMKXLTIPNLKEFPNLUYGAIKOGSTDSSRFVQELAMSRMOPHMKV 199

Qy 469 DLKMDNTRIHFQBDWKLITLTFIGNDLHYCENPAAALATEVQHQIQQALDILS 543
 Db 200 RMRMDPRVNMTSDWKLVLTFLFIGNN--DFCIDIYPPEPEKTVDWHRRNMLKTYRLD 256

Qy 544 EYPRAFVNLVTVLEIWLRL-----YQEKKWCPRMILSICPCVILKFDNME 593
 Db 257 NYPRMLNTIVPA---PNLRLPLNLGLPPICVS-----TLRFECPLM--GKGKGQ 302

Qy 594 LATLIEFNKKFOEKTQHOLIESGRYDTRBDFTVVQPRPENUDMPKTSBGLDPNSFPAPDC 653
 Db 303 LDYLEGMKRMKAQDFELANPREF-EPTINQVPSQDFPRRSQGDTRFSEDC 361

Qy 654 FHSKSKHSRAASALWNMMLSPVGQKT--TRHKFENKINTCPCNQVQPFATYKNS 707
 Db 362 FHSORGHAANSIWNMMLPGLGHSGGATHLFE--TFRCPSEMWPFLTRENS 414

RESULT 14

DE ABB9611
 ID ABB9611 standard; Protein; 148 AA.

XX AC ABB9611;

XX DT 21-JUN-2002 (first entry)

DE Human testicular antigen SEQ ID NO: 1495.

XX PN Human; testicular antigen; testes; cancer; metastasis; immune disorder;
 KW reproductive system disorder; urinary system disorder; gene therapy;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;

| | | |
|----|--|----------------------------------|
| KW | gastrointestinal disease; infection; cytostatic. | 25-SEP-2000; 2000US-0234998. |
| KX | | PR 26-SEP-2000; 2000US-0235084. |
| OS | Homo sapiens. | PR 27-SEP-2000; 2000US-0235034. |
| KX | | PR 27-SEP-2000; 2000US-0235036. |
| PN | WO200155317-A2. | PR 29-SEP-2000; 2000US-0236377. |
| XX | | PR 29-SEP-2000; 2000US-0236367. |
| PD | 02-AUG-2001. | PR 29-SEP-2000; 2000US-0236369. |
| PR | 17-JAN-2001; 2001WO-US01329. | PR 29-SEP-2000; 2000US-0236370. |
| XX | | PR 02-OCT-2000; 2000US-0236802. |
| PR | 31-JAN-2000; 2000US-0179065. | PR 02-OCT-2000; 2000US-0237037. |
| PR | 04-FEB-2000; 2000US-0180628. | PR 02-OCT-2000; 2000US-0237038. |
| PR | 24-FEB-2000; 2000US-0184664. | PR 02-OCT-2000; 2000US-0237039. |
| PR | 07-MAR-2000; 2000US-0186350. | PR 02-OCT-2000; 2000US-0237040. |
| PR | 16-MAR-2000; 2000US-0189874. | PR 13-OCT-2000; 2000US-0239935. |
| PR | 17-MAR-2000; 2000US-0190075. | PR 13-OCT-2000; 2000US-0239937. |
| PR | 18-APR-2000; 2000US-0198123. | PR 20-OCT-2000; 2000US-0240960. |
| PR | 19-MAY-2000; 2000US-0205515. | PR 20-OCT-2000; 2000US-0241221. |
| PR | 07-JUN-2000; 2000US-0209457. | PR 20-OCT-2000; 2000US-0241785. |
| PR | 26-JUN-2000; 2000US-0214886. | PR 20-OCT-2000; 2000US-0241786. |
| PR | 30-JUN-2000; 2000US-0215135. | PR 20-OCT-2000; 2000US-0241787. |
| PR | 07-JUL-2000; 2000US-0216647. | PR 20-OCT-2000; 2000US-0241808. |
| PR | 07-JUL-2000; 2000US-0216880. | PR 20-OCT-2000; 2000US-0241809. |
| PR | 11-JUL-2000; 2000US-0217487. | PR 20-OCT-2000; 2000US-0241936. |
| PR | 11-JUL-2000; 2000US-0217496. | PR 01-NOV-2000; 2000US-0241937. |
| PR | 14-JUL-2000; 2000US-0218290. | PR 08-NOV-2000; 2000US-02446414. |
| PR | 26-JUL-2000; 2000US-0220963. | PR 08-NOV-2000; 2000US-02446475. |
| PR | 26-JUL-2000; 2000US-0220964. | PR 08-NOV-2000; 2000US-02446476. |
| PR | 14-AUG-2000; 2000US-0224518. | PR 08-NOV-2000; 2000US-02446477. |
| PR | 14-AUG-2000; 2000US-0224519. | PR 08-NOV-2000; 2000US-02446478. |
| PR | 14-AUG-2000; 2000US-0225213. | PR 08-NOV-2000; 2000US-02446533. |
| PR | 14-AUG-2000; 2000US-0225214. | PR 08-NOV-2000; 2000US-02446534. |
| PR | 14-AUG-2000; 2000US-0225266. | PR 08-NOV-2000; 2000US-02446535. |
| PR | 14-AUG-2000; 2000US-0225267. | PR 08-NOV-2000; 2000US-02446536. |
| PR | 14-AUG-2000; 2000US-0225268. | PR 08-NOV-2000; 2000US-02446537. |
| PR | 14-AUG-2000; 2000US-0225270. | PR 08-NOV-2000; 2000US-02446538. |
| PR | 14-AUG-2000; 2000US-0225447. | PR 08-NOV-2000; 2000US-02446539. |
| PR | 14-AUG-2000; 2000US-0225757. | PR 08-NOV-2000; 2000US-02446609. |
| PR | 14-AUG-2000; 2000US-0225758. | PR 08-NOV-2000; 2000US-02446610. |
| PR | 14-AUG-2000; 2000US-0225759. | PR 08-NOV-2000; 2000US-02446611. |
| PR | 18-AUG-2000; 2000US-0226279. | PR 08-NOV-2000; 2000US-02446613. |
| PR | 22-AUG-2000; 2000US-0226681. | PR 08-NOV-2000; 2000US-02446707. |
| PR | 14-AUG-2000; 2000US-0226868. | PR 08-NOV-2000; 2000US-02446718. |
| PR | 22-AUG-2000; 2000US-0227182. | PR 08-NOV-2000; 2000US-02446729. |
| PR | 22-AUG-2000; 2000US-0227009. | PR 08-NOV-2000; 2000US-02446730. |
| PR | 30-AUG-2000; 2000US-0228924. | PR 08-NOV-2000; 2000US-02446741. |
| PR | 06-SEP-2000; 2000US-0229437. | PR 08-NOV-2000; 2000US-02446751. |
| PR | 01-SEP-2000; 2000US-0229343. | PR 08-NOV-2000; 2000US-02446761. |
| PR | 01-SEP-2000; 2000US-0229344. | PR 08-NOV-2000; 2000US-02446778. |
| PR | 01-SEP-2000; 2000US-0229345. | PR 08-NOV-2000; 2000US-02446789. |
| PR | 03-SEP-2000; 2000US-0229509. | PR 17-NOV-2000; 2000US-02449209. |
| PR | 05-SEP-2000; 2000US-0229513. | PR 17-NOV-2000; 2000US-02449210. |
| PR | 06-SEP-2000; 2000US-0230437. | PR 17-NOV-2000; 2000US-02449211. |
| PR | 06-SEP-2000; 2000US-0230438. | PR 17-NOV-2000; 2000US-02449212. |
| PR | 08-SEP-2000; 2000US-0231242. | PR 17-NOV-2000; 2000US-02449213. |
| PR | 08-SEP-2000; 2000US-0231243. | PR 17-NOV-2000; 2000US-02449214. |
| PR | 08-SEP-2000; 2000US-0231244. | PR 17-NOV-2000; 2000US-02449215. |
| PR | 08-SEP-2000; 2000US-0231413. | PR 17-NOV-2000; 2000US-02449216. |
| PR | 08-SEP-2000; 2000US-0231414. | PR 17-NOV-2000; 2000US-02449217. |
| PR | 08-SEP-2000; 2000US-0232080. | PR 17-NOV-2000; 2000US-02449218. |
| PR | 08-SEP-2000; 2000US-0232081. | PR 17-NOV-2000; 2000US-02449219. |
| PR | 14-SEP-2000; 2000US-0232401. | PR 17-NOV-2000; 2000US-02449220. |
| PR | 14-SEP-2000; 2000US-0232397. | PR 17-NOV-2000; 2000US-02449225. |
| PR | 14-SEP-2000; 2000US-0232398. | PR 17-NOV-2000; 2000US-0251988. |
| PR | 14-SEP-2000; 2000US-0232399. | PR 03-DEC-2000; 2000US-0256719. |
| PR | 14-SEP-2000; 2000US-0232423. | PR 06-DEC-2000; 2000US-0251479. |
| PR | 14-SEP-2000; 2000US-0232474. | PR 08-DEC-2000; 2000US-0251060. |
| PR | 14-SEP-2000; 2000US-0232474. | PR 01-DEC-2000; 2000US-0250311. |
| PR | 14-SEP-2000; 2000US-0233064. | PR 05-DEC-2000; 2000US-0251030. |
| PR | 14-SEP-2000; 2000US-0233065. | PR 05-DEC-2000; 2000US-0251889. |
| PR | 21-SEP-2000; 2000US-0234223. | PR 08-DEC-2000; 2000US-0251989. |
| PR | 21-SEP-2000; 2000US-0234274. | PR 08-DEC-2000; 2000US-0251990. |
| PR | 25-SEP-2000; 2000US-0234997. | PR 11-DEC-2000; 2000US-0254007. |
| PR | 05-JAN-2001; 2001US-0259678. | |

XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PT Roben CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483232/52.
XX
PT Nucleic acids encoding 973 human testicular antigen polypeptides, useful for preventing, diagnosing and/or treating testicular cancer -
XX
PS Claim 11; SEQ ID NO 1495; 766pp; English.
XX
CC The present invention provides the protein and coding sequences of 973 human testicular antigens, and fragments of their genomic sequences. The sequences can be used in the treatment of cardiovascular, urinary system, reproductive system, immune, respiratory, neurological and gastrointestinal disorders, infections, and particularly cancer, especially testicular cancers. The present sequence is a protein of the invention.
CC
XX Sequence 148 AA:

| | | | | |
|-----------------------|-------|--------------|---------|----------------|
| Query Match | 7.2% | Score 560; | DB 22; | Length 148; |
| Best Local Similarity | 77.9% | Pred. No. 4 | _ge-42; | |
| Matches | 113; | Conservative | 11; | Mismatches 17; |
| | | | | Indels 4; |
| | | | | Gaps 3; |

| | | | |
|----|------|--|------|
| OY | 1013 | RALWNTNMLPPLSKESTETDLRRAEMPTRCPTONBPFRLTPRNNTYIKAENWGSDFL | 1072 |
| Db | 4 | QSLWTNNMLPPLASKESTETDLRRAEMPTRCPTONBPFRLTPRNNTYIKAENWGSDFL | 63 |

| | | | |
|----|------|--|------|
| OY | 1073 | C'TEWKASNSVPSVHQLRPADTKWVAALGSDLTAVGARNPNSSDLSITSWRSQG-LSMSIGG | 1131 |
| Db | 64 | C'TEWKASNSVPSVHQLRPADTKWVAALGSDLTAVGARNPNSSSEPHILEGTLLGAGLEG | 123 |

| | | | |
|----|------|---------------------------|------|
| OY | 1132 | DGSNL-EHTTLPMN-LKKFNPVIL | 1153 |
| Db | 124 | MGNLGDHSHTAQHSEEVQPLPPWLL | 148 |

RESULT 15
ID AM95420
ID AM95420 standard; Protein: 148 AA.
XX
AC AM95420;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen SEQ ID NO: 4078.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PP 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-017905.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.

| | | |
|----|--------------|-----------------|
| PR | 11-JUL-2000; | 2000US-0217497. |
| PR | 11-JUL-2000; | 2000US-0217496. |
| PR | 14-JUL-2000; | 2000US-0218290. |
| PR | 26-JUL-2000; | 2000US-0220963. |
| PR | 26-JUL-2000; | 2000US-0220964. |
| PR | 14-AUG-2000; | 2000US-0224318. |
| PR | 14-AUG-2000; | 2000US-0224519. |
| PR | 14-AUG-2000; | 2000US-0225213. |
| PR | 14-AUG-2000; | 2000US-0225214. |
| PR | 14-AUG-2000; | 2000US-0225266. |
| PR | 14-AUG-2000; | 2000US-0225267. |
| PR | 14-AUG-2000; | 2000US-0225268. |
| PR | 14-AUG-2000; | 2000US-0225270. |
| PR | 14-AUG-2000; | 2000US-0225457. |
| PR | 14-AUG-2000; | 2000US-0225758. |
| PR | 14-AUG-2000; | 2000US-0225759. |
| PR | 22-AUG-2000; | 2000US-0226581. |
| PR | 22-AUG-2000; | 2000US-0226868. |
| PR | 23-AUG-2000; | 2000US-0227009. |
| PR | 30-AUG-2000; | 2000US-0228924. |
| PR | 01-SEP-2000; | 2000US-0229343. |
| PR | 01-SEP-2000; | 2000US-0229344. |
| PR | 01-SEP-2000; | 2000US-0229345. |
| PR | 05-SEP-2000; | 2000US-0229509. |
| PR | 05-SEP-2000; | 2000US-0229513. |
| PR | 06-SEP-2000; | 2000US-0230437. |
| PR | 06-SEP-2000; | 2000US-0230438. |
| PR | 08-SEP-2000; | 2000US-0231242. |
| PR | 08-SEP-2000; | 2000US-0231243. |
| PR | 08-SEP-2000; | 2000US-0231244. |
| PR | 08-SEP-2000; | 2000US-0231434. |
| PR | 08-SEP-2000; | 2000US-0231435. |
| PR | 08-SEP-2000; | 2000US-0232080. |
| PR | 08-SEP-2000; | 2000US-0232168. |
| PR | 12-SEP-2000; | 2000US-0232197. |
| PR | 14-SEP-2000; | 2000US-0232197. |
| PR | 14-SEP-2000; | 2000US-0232199. |
| PR | 14-SEP-2000; | 2000US-0232400. |
| PR | 14-SEP-2000; | 2000US-0232401. |
| PR | 14-SEP-2000; | 2000US-0233063. |
| PR | 14-SEP-2000; | 2000US-0233064. |
| PR | 14-SEP-2000; | 2000US-0233065. |
| PR | 21-SEP-2000; | 2000US-0234223. |
| PR | 21-SEP-2000; | 2000US-0234274. |
| PR | 25-SEP-2000; | 2000US-0234397. |
| PR | 25-SEP-2000; | 2000US-0234398. |
| PR | 26-SEP-2000; | 2000US-0235084. |
| PR | 27-SEP-2000; | 2000US-0235034. |
| PR | 27-SEP-2000; | 2000US-0235036. |
| PR | 29-SEP-2000; | 2000US-0236227. |
| PR | 29-SEP-2000; | 2000US-0236367. |
| PR | 29-SEP-2000; | 2000US-0236368. |
| PR | 29-SEP-2000; | 2000US-0236369. |
| PR | 29-SEP-2000; | 2000US-0236370. |
| PR | 02-OCT-2000; | 2000US-0236392. |
| PR | 02-OCT-2000; | 2000US-0237037. |
| PR | 02-OCT-2000; | 2000US-0237038. |
| PR | 02-OCT-2000; | 2000US-0237039. |
| PR | 02-OCT-2000; | 2000US-0237040. |
| PR | 13-OCT-2000; | 2000US-0239335. |
| PR | 13-OCT-2000; | 2000US-0239337. |
| PR | 20-OCT-2000; | 2000US-0240660. |
| PR | 20-OCT-2000; | 2000US-0241221. |
| PR | 20-OCT-2000; | 2000US-0241785. |
| PR | 20-OCT-2000; | 2000US-0241786. |
| PR | 20-OCT-2000; | 2000US-0241787. |
| PR | 20-OCT-2000; | 2000US-0241809. |

PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-024927.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249264.
 PR 05-DEC-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-024927.
 PR 17-NOV-2000; 2000US-0249299.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 05-DEC-2000; 2000US-0251989.
 PR 05-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Rosen CA, Barash SC, Ruben SM;
 XX DR WPI; 2001-465570/50.
 XX DR N-PSDB; AAL01390.
 PT Isolated nucleic acid molecule encoding a reproductive system antigen
 PT is used in preventing, treating or ameliorating a medical condition -
 XX Claim 11; SEQ ID NO 4078; 1297pp + Sequence Listing; English.
 CC The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a protein of the invention.
 XX SQ Sequence 148 AA;

Query Match Score 560; DB 22; Length 148;
 Best Local Similarity 77.9%; Prd. No. 4.9e-42;
 Matches 113; Conservative 11; Mismatches 17; Indels 4; Gaps 3;

PD XX
 PR XX
 22-JAN-2002; 2002WO-US01715.
 PR XX
 24-JAN-2001; 2001WS-264049P.
 PA XX
 (LEXI-) LEXICON GENETICS INC.
 PI XX
 Yu X., Miranda M., Turner CA;
 DR XX
 WPI; 2002-599797/64.
 P-PSSB; ABB09555.
 PT XX
 Polynucleotides encoding human lipases that are structurally related to
 animal lipases, particularly phospholipase B, useful for drug
 screening, diagnosis and in gene therapy of biological disorders - .
 PS XX
 Claim 1; Page 36-37; 44pp; English.
 The invention relates to a novel human lipase (NHL; ABB09555, ABB09556)
 and to nucleic acids encoding it (ABQ77623, ABQ77624). The NHL has
 structural similarity with animal lipases, particularly phospholipase B.
 Polynucleotides encoding NHL were obtained using human genomic sequences
 in conjunction with human thyroid and brain cDNAs. The NHL gene is
 located on chromosome 2, and contains a C/T polymorphism at position
 3953 of the open reading frame (ORF), resulting in an Ala/Val
 substitution at position 1318 in the protein. NHL nucleotides and
 proteins are useful for treating disorders such as inflammatory or
 proliferative disease, infectious disease, clotting disorders, and
 cancer. They can also be used in screening for compounds useful in
 the treatment of mental, biological or medical disorders, as diagnostic
 reagents, in clinical trial monitoring and in cosmetic and nutraceutical
 applications. NHL nucleotides can additionally be used in the detection
 of disease-associated mutations, in the analysis of gene expression, for
 the recombinant expression of NHL to generate transgenic animals, in
 gene therapy, and as part of ribozyme and/or triple helix sequences
 useful in the modulation of NHL gene expression. The present sequence
 represents cDNA encoding the Ala 1318 variant of NHL.
 Sequence 4377 BP; 1085 A; 1217 C; 1155 G; 920 T; 0 other;
 Query Match 100.0%; Score 4377; DB 24; Length 4377;
 Best Local Similarity 100.0%; Precl. No. 0;
 Matches 4377; Conservative 0; Mismatches 0; Indels 0; Gaps 0
 OY 1 ATGGGCGCTGGGCCGGCATTTCCCTCTGGCTCTGCGCTCTGGGGCAAGGACC 60
 Db 1 ATGGGCGCTGGGCCGGCATTTCCCTCTGGCTCTGCGCTCTGGGGCAAGGACC 60
 OY 61 CCTCGATCCATACCTCTCTAGAAGACTACATTGAAAGGGCAGCTATGCCAGAGACC 120
 Db 61 CCTCGATCCATACCTCTCTAGAAGAGTACATTGAAAGGGCAGCTATGCCAGAGACC 120
 OY 121 CTGAGAAGTCCTCCATCCATGCAACCAATAATTAGGGATGATATGCCCTCAA 180
 Db 121 CTGAGAAGTCCTCCATCCATGCAACCAATAATTAGGGATGATATGCCCTCAA 180
 OY 181 TCACTTCACTCTCTAGGCCTCTATTAATTAGGGATGAGCAGCTCTTCAGACATCATCATCAGATAATTAGGGATGATATGCCCTCAA 240
 Db 181 TCACTTCACTCTCTAGGCCTCTATTAATTAGGGATGAGCAGCTCTTCAGACATCATCATCAGATAATTAGGGATGATATGCCCTCAA 240
 OY 241 ATTCCTCCAGACCCGGGAGCGGCATCTGGAGAGGAAGCTGGACTGAAGGGCACAG 300
 Db 241 ATTCCTCCAGACCCGGGAGCGGCATCTGGAGAGGAAGCTGGACTGAAGGGCACAG 300
 OY 301 CAGGTGTGCAATGGGAGTGACAGACTCTTCAGACATCATCAGATAATTAGGGATGATATGCCCTCAA 360
 Db 301 CAGGTGTGCAATGGGAGTGACAGACTCTTCAGACATCATCAGATAATTAGGGATGATATGCCCTCAA 360
 361 GTTCCCAATGCTGTGCGCACACTGAGAAGAGGAGCTACCCACAGTGGCTGAGAC 420
 Db 361 GTTCCCAATGCTGTGCGCACACTGAGAAGAGGAGCTACCCACAGTGGCTGAGAC 420

| | | | | |
|----|------|--|------|----|
| QY | 1501 | GAAGCTGGAGATAATAACCTGTATTAGGCGCCAATGACCTTGATTCGCAAT | 1560 | Db |
| Db | 1502 | GATCGTGGCACTATCTCCAGACTCACAGAACATGGAAAGGCCTGGCACAT | 1561 | Qy |
| Qy | 1561 | GATCGTGGCACTATCTCCAGACTCACAGAACATGGAAAGGCCTGGCACAT | 1560 | Db |
| Db | 1561 | GATCGTGGCACTATCTCCAGACTCACAGAACATGGAAAGGCCTGGCACAT | 1620 | Qy |
| Qy | 1620 | GATCGTGGCACTATCTCCAGACTCACAGAACATGGAAAGGCCTGGCACAT | 1620 | Db |
| Db | 1621 | CTCCATGCTGAGGTCTCGGGCATTTGTRACCTGGTCAAC | 1680 | Qy |
| Qy | 1680 | CTCCATGCTGAGGTCTCGGGCATTTGTRACCTGGTCAAC | 1680 | Db |
| Db | 1681 | CTGAGGGAGCTGACAGGAAAAAAGTACTGCCAGGATGATCTCAGCTCTCG | 1740 | Qy |
| Qy | 1740 | CTGAGGGAGCTGACAGGAAAAAAGTACTGCCAGGATGATCTCAGCTCTCG | 1740 | Db |
| Db | 1741 | TGTCCCTGCTCTCAGTTGATTAACCTCAAGAGAACTGCTACCCATGGAAATT | 1800 | Qy |
| Qy | 1800 | TGTCCCTGCTCTCAGTTGATTAACCTCAAGAGAACTGCTACCCATGGAAATT | 1800 | Db |
| Db | 1801 | AACAGAAGTTCAAGAGAAGACCCACACTGAGATGAGTGGCATGACMAAG | 1860 | Qy |
| Qy | 1860 | AACAGAAGTTCAAGAGAAGACCCACACTGAGATGAGTGGCATGACMAAG | 1860 | Db |
| Db | 1861 | GAAGTTTACTGCTGTTGACCGCTTGAIAACTGAGATGATCTCAGCTCTCG | 1920 | Qy |
| Qy | 1920 | GAAGTTTACTGCTGTTGACCGCTTGAIAACTGAGATGATCTCAGCTCTCG | 1920 | Db |
| Db | 1861 | GAAGTTTACTGCTGTTGACCGCTTGAIAACTGAGATGATCTCAGCTCTCG | 1920 | Qy |
| Qy | 1921 | GAAGGATGCTGAGAACTTCTCTCGCTCTGACTGTTCCACTTCAGCAGACT | 1980 | Db |
| Db | 1980 | GAAGGATGCTGAGAACTTCTCTCGCTCTGACTGTTCCACTTCAGCAGACT | 1980 | Qy |
| Qy | 1980 | GAAGGATGCTGAGAACTTCTCTCGCTCTGACTGTTCCACTTCAGCAGACT | 1980 | Db |
| Db | 1981 | CACTCCGAGCAGCCAGTGTCTCTGACAAATATGCTGGAGCTTCTGAGAGAC | 2040 | Qy |
| Qy | 2040 | CACTCCGAGCAGCCAGTGTCTCTGACAAATATGCTGGAGCTTCTGAGAGAC | 2040 | Db |
| Db | 2041 | ACTCTCATAGTTGAAACAGTCATATCAGTCGAGCAGCTCCAGCGTT | 2100 | Qy |
| Qy | 2100 | ACTCTCATAGTTGAAACAGTCATATCAGTCGAGCAGCTCCAGCGTT | 2100 | Db |
| Db | 2041 | ACTCTCATAGTTGAAACAGTCATATCAGTCGAGCAGCTCCAGCGTT | 2100 | Qy |
| Qy | 2100 | ACTCTCATAGTTGAAACAGTCATATCAGTCGAGCAGCTCCAGCGTT | 2100 | Db |
| Db | 2101 | CTGAGGACTTACAAGAACAGCATGGAGGTATGGACACTGGCTGGCTCAT | 2160 | Qy |
| Qy | 2160 | CTGAGGACTTACAAGAACAGCATGGAGGTATGGACACTGGCTGGCTCAT | 2160 | Db |
| Db | 2161 | GCCCCTTCCTCCACCTTACCTTACCTTACCTTACCTTACCTTACCTT | 2220 | Qy |
| Qy | 2220 | GCCCCTTCCTCCACCTTACCTTACCTTACCTTACCTTACCTTACCTT | 2220 | Db |
| Db | 2221 | GTGGTGTCTGGGGAGTTCTGAGCGCTGGCAATGGAAATGGCTCCAAACAGCAGCA | 2280 | Qy |
| Qy | 2280 | GTGGTGTCTGGGGAGTTCTGAGCGCTGGCAATGGAAATGGCTCCAAACAGCAGCA | 2280 | Db |
| Db | 2221 | GTGGTGTCTGGGGAGTTCTGAGCGCTGGCAATGGAAATGGCTCCAAACAGCAGCA | 2280 | Qy |
| Qy | 2280 | GTGGTGTCTGGGGAGTTCTGAGCGCTGGCAATGGAAATGGCTCCAAACAGCAGCA | 2280 | Db |
| Db | 2281 | CTCCCGATCTCACACACAGTATGGGAGTGTATCACAGTCGAGGGGGCTCC | 2340 | Qy |
| Qy | 2340 | CTCCCGATCTCACACACAGTATGGGAGTGTATCACAGTCGAGGGGGCTCC | 2340 | Db |
| Db | 2341 | CTGGAGATGTCGACCTTACCTTCTCGAGGTTAACAGAACCTCAGCC | 2400 | Qy |
| Qy | 2400 | CTGGAGATGTCGACCTTACCTTCTCGAGGTTAACAGAACCTCAGCC | 2400 | Db |
| Db | 2401 | TACGGCTGGCACGGGTGATGCCATGACAGGAATCTCTCAATCAAGCTTCCC | 2460 | Qy |
| Qy | 2460 | TACGGCTGGCACGGGTGATGCCATGACAGGAATCTCTCAATCAAGCTTCCC | 2460 | Db |
| Db | 2461 | GGAGGAAGCTGAGATCTTACCTGAGCACTGAGATGAGATGAGATGAGAT | 2520 | Qy |
| Qy | 2520 | GGAGGAAGCTGAGATCTTACCTGAGCACTGAGATGAGATGAGATGAGAT | 2520 | Db |
| Db | 2521 | GATCTAGAGTAATTCATGAGACTGGAGGCTGAGCTGAGCTGAGGAGC | 2580 | Qy |
| Qy | 2580 | GATCTAGAGTAATTCATGAGACTGGAGGCTGAGCTGAGGAGC | 2580 | Db |
| Db | 2581 | GATTAGTAGTACTGAGATTCAGGAAATGTCAGCACTCCACGGCCTGACACTTGTGACCAT | 2640 | Qy |
| Qy | 2640 | GATTAGTAGTACTGAGATTCAGGAAATGTCAGCACTCCACGGCCTGACACTTGTGACCAT | 2640 | Db |
| Db | 2641 | CTCGCAATGCTGACTCTCTCATAGAGATGCGCTGCTGCTGCTGCTG | 2700 | Qy |
| Qy | 2700 | CTCGCAATGCTGACTCTCTCATAGAGATGCGCTGCTGCTGCTG | 2700 | Db |
| Db | 2701 | ACTCTGAACTGACATCTGGGGGGTGTCTCGAGGAACTCTGAGATGCTCTCG | 2760 | Qy |
| Qy | 2760 | ACTCTGAACTGACATCTGGGGGGTGTCTCGAGGAACTCTGAGATGCTCTCG | 2760 | Db |
| Db | 2761 | GTGAGCAGGCCAGGCTATCGGGGGTGTCTCGAGGAACTCTGAGATGCTCTCG | 2820 | Qy |
| Qy | 2820 | GTGAGCAGGCCAGGCTATCGGGGGTGTCTCGAGGAACTCTGAGATGCTCTCG | 2820 | Db |
| Db | 2821 | CTACCCAGGCTGGAGGCTCTACCGGACGATGACGAGCTGCGGAGCTGCGG | 2880 | Qy |
| Qy | 2880 | CTACCCAGGCTGGAGGCTCTACCGGACGATGACGAGCTGCGGAGCTGCGG | 2880 | Db |
| Db | 2881 | TCAGCCGTTGACCGAGGAGACTCTCTGGGGCTGAGGAGACTTCCAGACAG | 2940 | Qy |
| Qy | 2940 | TCAGCCGTTGACCGAGGAGACTCTCTGGGGCTGAGGAGACTTCCAGACAG | 2940 | Db |
| Db | 2881 | TCAGCCGTTGACCGAGGAGACTCTCTGGGGCTGAGGAGACTTCCAGACAG | 2940 | Qy |
| Qy | 2941 | ATCCAGCTCTGCTCTGGAGGATEGGCTCCAGATACTGCTCTGCTGCTG | 3000 | Db |
| Db | 3000 | ATCCACCCAAATCAGAAATTCACCTCCAGCTCCAGGAGCTTCCAGACAG | 3000 | Qy |
| Qy | 3000 | ATCCACCCAAATCAGAAATTCACCTCCAGCTCCAGGAGCTTCCAGACAG | 3000 | Db |
| Db | 3061 | GAACCACTTGAGCAAACAGAGACCCCTGGCCAGGCCCCTTGGACCAATATGCT | 3120 | Qy |
| Qy | 3120 | GAACCACTTGAGCAAACAGAGACCCCTGGCCAGGCCCCTTGGACCAATATGCT | 3120 | Db |
| Db | 3061 | GAACCACTTGAGCAAACAGAGACCCCTGGCCAGGAGATGCGCATCCATCGT | 3120 | Qy |
| Qy | 3120 | GAACCACTTGAGCAAACAGAGACCCCTGGCCAGGAGATGCGCATCCATCGT | 3120 | Db |
| Db | 3121 | CCCACTCAGATACTGAGCCCTCTGAGAACCCCTGGAAATAGTACTACAGTACCCATC | 3180 | Qy |
| Qy | 3180 | CCCACTCAGATACTGAGCCCTCTGAGAACCCCTGGAAATAGTACTACAGTACCCATC | 3180 | Db |
| Db | 3181 | AAGCCAGCATGAGACTGGGGAGTGTCTCTGTGAGCTGGAGGAGCTCCAT | 3240 | Qy |
| Qy | 3240 | AAGCCAGCATGAGACTGGGGAGTGTCTCTGTGAGCTGGAGGAGCTCCAT | 3240 | Db |
| Db | 3241 | AGTGTCTCACCTCTGTCACCTCCACCTCCACCTCCACCTCCACCTCC | 3300 | Qy |
| Qy | 3300 | AGTGTCTCACCTCTGTCACCTCCACCTCCACCTCCACCTCCACCTCC | 3300 | Db |
| Db | 3301 | GTTGACTCTCTGACTACAGCACTGGAGGTGTCGACCAACAACTCAGTACCTACCA | 3360 | Qy |
| Qy | 3360 | GTTGACTCTCTGACTACAGCACTGGAGGTGTCGACCAACAACTCAGTACCTACCA | 3360 | Db |
| Db | 3301 | GTTGACTCTCTGACTACAGCACTGGAGGTGTCGACCAACAACTCAGTACCTACCA | 3360 | Qy |
| Qy | 3360 | GTTGAGGGACTCTCTGGAGGTGGGGAGATGGAGACTCACCCACA | 3420 | Db |
| Db | 3361 | TCTGGAGGGACTCTCTGGAGGTGGGGAGATGGAGACTCACCCACA | 3420 | Qy |
| Qy | 3420 | TCTGGAGGGACTCTCTGGAGGTGGGGAGATGGAGACTCACCCACA | 3420 | Db |
| Db | 3421 | CTGCCAACATCTGAGAGCTTACCTTACCTCTGGCTCTCTGACACTCG | 3480 | Qy |
| Qy | 3480 | CTGCCAACATCTGAGAGCTTACCTTACCTCTGGCTCTCTGACACTCG | 3480 | Db |
| Db | 3481 | GAGGGAGCAGCAGGACTAAATGTGCAAGGGAGGGGGAGACTAGCTGCCAGCC | 3540 | Qy |
| Qy | 3540 | GAGGGAGCAGCAGGACTAAATGTGCAAGGGAGGGGGAGACTAGCTGCCAGCC | 3540 | Db |
| Db | 3541 | CAGGCTGGGACCTCTGGAGGAACTGAGGAGGAGGAGACTGGAGACTCACCTCT | 3600 | Qy |
| Qy | 3600 | CAGGCTGGGACCTCTGGAGGAGGAGGAGACTGGAGACTCACCTCT | 3600 | Db |
| Db | 3601 | TGGAGCTGCTGACTCTCTGGGGCACTCTGGGGAGGGGGAGACTAGCTGCCAGCC | 3660 | Qy |
| Qy | 3660 | TGGAGCTGCTGACTCTCTGGGGCACTCTGGGGAGGGGGAGACTAGCTGCCAGCC | 3660 | Db |
| Db | 3661 | GGAGCCACTTGGCACGGAATGTCAGCACTCCACGGCCCTGACACTCTCT | 3720 | Qy |

/note= "Single nucleotide polymorphism (SNP); leads to a Val to Ala substitution at position 1318 of the protein"

Db 3661 GAGGCCCACTTGGCCACGGATATGTTCAGCACATCCACAGGCCCCCTGGACATCCTCTCT 3720
 Qy 3721 GAGGAGCTCCCAGGGTTCTGAAACGGTGCATGAGCGTGCTGCTACCTGTCAC 3780
 Db 3721 GAGGAGCTCCCAGGGTTCTGAAACGGTGCATGAGCGTGCTGCTACCTGTCAC 3780
 Qy 3781 CAGGGCAAGGCGGAATGTGCAATGCTGGAGCTAGAACACTGACTTCCTCAGA 3840
 Db 3781 CAGGGCAAGGCGGAATGTGCAATGCTGGAGCTAGAACACTGACTTCCTCAGA 3840
 Qy 3841 CACTCGCAAGCTCCCTGAGAAGCAAGACTGAGAAGTGAAGTGAACCTCAGAT 3900
 Db 3841 CACTCGCAAGCTCCCTGAGAAGCAAGACTGAGAAGTGAAGTGAACCTCAGAT 3900
 Qy 3901 GSCATTCAGTCTCCGAGGACTGTTTCACTCTCAGACAGCGAGGTTTCTGGT 3960
 Db 3901 GSCATTCAGTCTCCGAGGACTGTTTCACTCTCAGACAGCGAGGTTTCTGGT 3960
 Qy 3960 GCTGCAAGCTCCCTGAGAAGCAAGACTGAGAAGTGAAGTGAACCTCAGAT 3960
 Db 3960 GCTGCAAGCTCCCTGAGAAGCAAGACTGAGAAGTGAAGTGAACCTCAGAT 3960
 Qy 3961 GTGCAAGCTCCCTGAGAAGCAAGACTGAGAAGTGAAGTGAACCTCAGAT 3960
 Db 3961 GTGCAAGCTCCCTGAGAAGCAAGACTGAGAAGTGAAGTGAACCTCAGAT 3960
 Qy 4021 ACCTCTCTCTCCGAGGTTTCACTCTCAGACCGGGCATGGCTGACCTGCATC 4080
 Db 4021 ACCTCTCTCTCCGAGGTTTCACTCTCAGACCGGGCATGGCTGACCTGCATC 4080
 Qy 4081 GCACTCTGAAACACATGCTGGACCCTGAGGACTTCCACACTTCACC 4140
 Db 4081 GCACTCTGAAACACATGCTGGACCCTGAGGACTTCCACACTTCACC 4140
 Qy 4141 CACAGCGAGGCAACTCAAGGCGCCCTGAGGCGCTTACCTACCCCTCGG 4200
 Db 4141 CACAGCGAGGCAACTCAAGGCGCCCTGAGGCGCTTACCTACCCCTCGG 4200
 Qy 4201 ARGAGCGATTGCTCCAGACAGCGCTGAGAAGGCCCGGAGGTCTACTGGCTC 4260
 Db 4201 ARGAGCGATTGCTCCAGACAGCGCTGAGAAGGCCCGGAGGTCTACTGGCTC 4260
 Qy 4261 CTAATGGCGAGGGAGTGGCTGTTGGGATCATGGGACAGTGTCTGGAGTC 4320
 Db 4261 CTAATGGCGAGGGAGTGGCTGTTGGGATCATGGGACAGTGTCTGGAGTC 4320
 Qy 4321 AGGAGAGTGGGGAGGAGAATCTCCAATGAGCCCTGGACTGGCCCTAG 4377
 Db 4321 AGGAGAGTGGGGAGGAGAATCTCCAATGAGCCCTGGACTGGCCCTAG 4377
 OS Homo sapiens.

RESULT 2

ID ABQ77624 Standard; cdna; 4377 BP.
 XX ABQ77624;
 AC
 DT 21-OCT-2002 (first entry)
 XX Human lipase NHL (Val 1318 variant) cdna.
 DE Human lipase NHL (Val 1318 variant) cdna.

Query Match 100.0%; Score 4375.4; DB 24; Length 4377;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4376; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGGTGGCCCAAGCAATTTCCTCTGGAGCTCTGGCTCTTGCGCAAGGCC 60
 Db 1 ATGGGGTGGCCCAAGCAATTTCCTCTGGAGCTCTGGCTCTTGCGCAAGGCC 60
 Qy 61 CCTCAGATCCATACCTCTCTAGAAAGTAGACATGGAGGCGAGCTATGGCGAGAGCC 120
 Db 61 CCTCAGATCCATACCTCTCTAGAAAGTAGACATGGAGGCGAGCTATGGCGAGAGCC 120
 Qy 121 CTGAGAGATTCATCCATGCAACCAAATAATTAGGAGTGAATATGCGCAATCTTAA 180
 Db 121 CTGAGAGATTCATCCATGCAACCAAATAATTAGGAGTGAATATGCGCAATCTTAA 180
 Qy 181 TCACTCTCTGAGGCTCTGATATAATTGCGAGGCAATTGCGAA 240
 Db 181 TCACTCTCTGAGGCTCTGATATAATTGCGAGGCAATTGCGAA 240
 Qy 241 ATTCTCCAGACCCAGGGAGGGAGTGGAGAAGCAAGCTGGACTGAAGGCCAG 300
 Db 241 ATTCTCCAGACCCAGGGAGGGAGTGGAGAAGCAAGCTGGACTGAAGGCCAG 300

Key Location/Qualifiers
 FT 1..4377
 FT /*tag= a
 FT /product= "Human lipase NHL"
 variation replace (3953, C)
 FT /*tag= b

| | | | | | | | |
|----|------|---|------|----|------|--|------|
| QY | 301 | CAGGTGTCATGGGAGTGTACAGTCCTTCAGACATCATCGATATTCTACTCTCT | 360 | QY | 1381 | GAACCACTCTTAATGCCCTCTAACCCAGCTGGCAGGAGCCGCGCTGAGGATCTA | 1440 |
| Db | 301 | CAGGTGTCATGGGAGTGTACAGTCCTTCAGACATCATCGATATTCTACTCTCT | 360 | Do | 1381 | GAAACACTCTTAATGCCCTCTAACCCAGCTGGCAGGAGCCGCGCTGAGGATCTA | 1440 |
| QY | 361 | GTTCGAATGCCTCTGTCACACTTGAAAGAGAGTCATACCCACGATGGTCTGAGAC | 420 | QY | 1441 | CCTGTCAGGCCAGGAGGCTGGTGTGACATGAGAAGATGACAGAGGATACACTTAG | 1500 |
| Db | 361 | GTTCGAATGCCTCTGTCACACTTGAAAGAGAGTCATACCCACGATGGTCTGAGAC | 420 | Do | 1441 | CCTGTCAGGCCAGGAGGCTGGTGTGACATGAGAAGATGACAGAGGATACACTTAG | 1500 |
| QY | 421 | TCTGGGATTCAGGCTCAAGAACCTGGTGAAGAACATGAAGAAGAACCTGCAACTGACTT | 480 | QY | 1501 | GAAACTGAAAGATAATACCCCTGTTAAGGGCAATGACCTCTGATGATTCTGAAAT | 1560 |
| Db | 421 | TCTGGGATTCAGGCTCAAGAACCTGGTGAAGAACATGAAGAAGAACCTGCAACTGACTT | 480 | Do | 1501 | GAAACTGAAAGATAATACCCCTGTTAAGGGCAATGACCTCTGATGATTCTGAAAT | 1560 |
| QY | 481 | CAATTGACTGAGACTCATCAAATGTTCTTCAGTAATGCAAGCCACTGTACCTGTC | 540 | QY | 1561 | GATCTGGTCAACTATTCTCCCAGAACTTACACAGACATGGTGTGAGATGTCAC | 1620 |
| Db | 481 | CAATTGACTGAGACTCATCAAATGTTCTTCAGTAATGCAAGCCACTGTACCTGTC | 540 | Do | 1561 | GATCTGGTCAACTATTCTCCCAGAACTTACACAGACATGGTGTGAGATGTCAC | 1620 |
| QY | 541 | CCCTCTGTCACAAGAACATGGCTGGCGGGGGCGTGTGAGAAGAACCTGCAACTGACTT | 600 | QY | 1621 | CTTCATGTTGAGGTCTCTCGGECATTGTTGAAACCTGGTCTCTGAGTT | 1680 |
| Db | 541 | CCCTCTGTCACAAGAACATGGCTGGCGGGGGCGTGTGAGAAGAACCTGCAACTGACTT | 600 | Do | 1621 | CTTCATGTTGAGGTCTCTCGGECATTGTTGAAACCTGGTCTCTGAGTT | 1680 |
| QY | 601 | GRCTACTCTGAGGAGTCCCGAGAGATTTAACCTGGTCTCTGAGTT | 660 | QY | 1680 | CTCAGGGAGCTTACAGAGGAAAAAAGTCACTGCCCCAGGATGTCAGTCTG | 1740 |
| Db | 601 | GRCTACTCTGAGGAGTCCCGAGAGATTTAACCTGGTCTCTGAGTT | 660 | Do | 1680 | CTCAGGGAGCTTACAGAGGAAAAAAGTCACTGCCCCAGGATGTCAGTCTG | 1740 |
| QY | 661 | GCAGAGGGCTCTGTCAGTATCAGGACTTCTGAGCTTCTGGCTTAACCTGTTAAT | 720 | QY | 1741 | TGTCCTCTGCTGAAAGTTGATTAACCTGACAGCTCTCCTCTGAAAT | 1800 |
| Db | 661 | GCAGAGGGCTCTGTCAGTATCAGGACTTCTGAGCTTCTGGCTTAACCTGTTAAT | 720 | Do | 1741 | TGTCCTCTGCTGAAAGTTGATTAACCTGACAGCTCTCCTCTGAAAT | 1800 |
| QY | 721 | TCTCTGAGGAGAACCCGGCTGGCCAGGACTCTGAGGAGGAGTCTGAGGTT | 780 | QY | 1801 | AACAGAAGATTTCAGGAGAAGACCACAACTGATTGAGAGTGGCGATATGACAGG | 1860 |
| Db | 721 | TCTCTGAGGAGAACCCGGCTGGCCAGGACTCTGAGGAGGAGTCTGAGGTT | 780 | Do | 1801 | AACAGAAGATTTCAGGAGAAGACCACAACTGATTGAGAGTGGCGATATGACAGG | 1860 |
| QY | 781 | TGAAACAGCCTCTGGCTCCCGAGGACTCTGAGGAGGAGTCTGAGGTT | 840 | QY | 1860 | 1861 GAAAGATTACTGTGTTGTCAGGAGGTTCTTGAAGAAGTCCAGGAGCTCG | 1920 |
| Db | 781 | TGAAACAGCCTCTGGCTCCCGAGGACTCTGAGGAGGAGTCTGAGGTT | 840 | Do | 1860 | 1861 GAAAGATTACTGTGTTGTCAGGAGGTTCTTGAAGAAGTCCAGGAGCTCG | 1920 |
| QY | 841 | CAACCTCTCTATGAGACCCATCTCACCTGGAGAACCCCGAGCTCCAGAT | 900 | QY | 1921 | GAAGGATTGCTCTACAACTCTCTCCCTCTGACTCTTCACTTCAGGACAGTCT | 1980 |
| Db | 841 | CAACCTCTCTATGAGACCCATCTCACCTGGAGAACCCCGAGCTCCAGAT | 900 | Do | 1921 | GAAGGATTGCTCTACAACTCTCTCCCTCTGACTCTTCACTTCAGGACAGTCT | 1980 |
| QY | 901 | TCTTACACGGCTCTGGCATCTCTGGATAAGATGATGAGGAGGAGTCTGAGGTT | 960 | QY | 1981 | CACTCCGGAGCAGCCAGTGTCTCTGAGAACAAATATGCTGGAGCTGTCGAG | 2040 |
| Db | 901 | TCTTACACGGCTCTGGCATCTCTGGATAAGATGATGAGGAGGAGTCTGAGGTT | 960 | Do | 1981 | CACTCCGGAGCAGCCAGTGTCTCTGAGAACAAATATGCTGGAGCTGTCGAG | 2040 |
| QY | 961 | GGCCATCTGAGTTAACACGGGAGGCAATGAGTCTCTCTCAGGAGCCCTAT | 1020 | QY | 2041 | ACTGTCCTAAGTTGAAACAGATCATATCACATGTCGAGGCTCCAGCTT | 2100 |
| Db | 961 | GGCCATCTGAGTTAACACGGGAGGCAATGAGTCTCTCTCAGGAGCCCTAT | 1020 | Do | 2041 | ACTGTCCTAAGTTGAAACAGATCATATCACATGTCGAGGCTCCAGCTT | 2100 |
| QY | 1021 | CTGTTCACTCTACAGAAACGCAACTACTTGACGAGCTGAGAAACCCAACTT | 1080 | QY | 2101 | CTGAGGACCTACAGAAACGCAACTACTTGACGAGCTGAGAAACCCAACTT | 2160 |
| Db | 1021 | CTGTTCACTCTACAGAAACGCAACTACTTGACGAGCTGAGAAACCCAACTT | 1080 | Do | 2101 | CTGAGGACCTACAGAAACGCAACTACTTGACGAGCTGAGAAACCCAACTT | 2160 |
| QY | 1081 | GGGTGAAGAGAGGAGGAGGAAATCAGATGTCCTGAGAACAGCCCTCGATGGTTCC | 1140 | QY | 2161 | GGCCCTCTGCTCTGACCCCTACCTCTGAGCTGAGAACATCCAAATT | 2220 |
| Db | 1081 | GGGTGAAGAGAGGAGGAGGAAATCAGATGTCCTGAGAACAGCCCTCGATGGTTCC | 1140 | Do | 2161 | GGCCCTCTGCTCTGACCCCTACCTCTGAGCTGAGAACATCCAAATT | 2220 |
| QY | 1141 | ACCTCACTGTCAGGCTGAGGCGCGTCACTACAGTAATGGAGCCCTGGTGTACT | 1200 | QY | 2221 | GGCTCTCTGGGGATCTCTGACCCGCTGGCAATGGATTGTCCTCAACACGAC | 2280 |
| Db | 1141 | ACCTCACTGTCAGGCTGAGGCGCGTCACTACAGTAATGGAGCCCTGGTGTACT | 1200 | Do | 2221 | GGCTCTCTGGGGATCTCTGACCCGCTGGCAATGGATTGTCCTCAACACGAC | 2280 |
| QY | 1201 | TACCGGAGGCTCTGGCTGGCGAGATGAGACATGGCACGGTTCACCTG | 1260 | QY | 2281 | CTCCCGGAGTGTGACCCACTTACCTTAATTCCTGGGGATTACAGAACCTCACAGC | 2340 |
| Db | 1201 | TACCGGAGGCTCTGGCTGGCGAGATGAGACATGGCACGGTTCACCTG | 1260 | Do | 2281 | CTCCCGGAGTGTGACCCACTTACCTTAATTCCTGGGGATTACAGAACCTCACAGC | 2340 |
| QY | 1261 | TACCGGAGGCTCTGGCTGGCGAGATGAGACATGGCACGGTTCACCTG | 1320 | QY | 2341 | CTGGAGGAGTGTGACCCACTTACCTTAATTCCTGGGGATTACAGAACCTCACAGC | 2400 |
| Db | 1261 | TACCGGAGGCTCTGGCTGGCGAGATGAGACATGGCACGGTTCACCTG | 1320 | Do | 2341 | CTGGAGGAGTGTGACCCACTTACCTTAATTCCTGGGGATTACAGAACCTCACAGC | 2400 |
| QY | 1321 | GGAAACATCTCTGGGAATTCACTCTGGCTGGAGCTGAGCTGGAA | 1380 | QY | 2401 | TACGGCTGGAGGAGTGTGACCCACTTACCTTAATTCCTGGGGATTACAGAACCTCACAGC | 2460 |
| Db | 1321 | GGAAACATCTCTGGGAATTCACTCTGGCTGGAGCTGAGCTGGAA | 1380 | Do | 2401 | TACGGCTGGAGGAGTGTGACCCACTTACCTTAATTCCTGGGGATTACAGAACCTCACAGC | 2460 |
| QY | 2461 | GGACAAAGGCTGAGGACTTATGAGCCAAGTCCAAACTCTGATGAGGATGAAGAT | 2520 | | | | |

| | | |
|----|----|--|
| XX | FH | Location/Qualifiers |
| CC | FT | 9..4385 |
| CC | FT | /*tag= a |
| CC | FT | /product = "Human LIPAM-1 protein" |
| CC | FT | 9..65 |
| CC | FT | /*tag= b |
| CC | FT | 66..4382 |
| CC | FT | /*tag= c |
| CC | FT | /product = "Human mature LIPAM-1 protein" |
| CC | FT | 9..71 |
| CC | FT | /*tag= d |
| CC | FT | 72..4382 |
| CC | FT | /*tag= e |
| CC | FT | /product = "Human mature LIPAM-1 protein" |
| CC | FT | 9..74 |
| CC | FT | /*tag= f |
| CC | FT | 75..4382 |
| CC | FT | /*tag= g |
| CC | FT | /product = "Human mature LIPAM-1 protein" |
| CC | FT | 9..77 |
| CC | FT | /*tag= h |
| CC | FT | 78..4382 |
| CC | FT | /*tag= i |
| CC | FT | /product = "Human mature LIPAM-1 protein" |
| CC | FT | 9..89 |
| CC | FT | /*tag= j |
| CC | FT | 90..4382 |
| CC | FT | /*tag= k |
| CC | FT | /product = "Human mature LIPAM-1 protein" |
| XX | RN | WO200394988-A2. |
| XX | PD | 28-NOV-2002. |
| XX | PP | 17-MAY-2002; 2002WO-US15688. |
| XX | PR | 18-MAY-2001; 2001US-292242P. |
| XX | PR | 25-MAY-2001; 2001US-29376P. |
| XX | PR | 01-JUN-2001; 2001US-29536P. |
| XX | PR | 06-JUL-2001; 2001US-30344P. |
| XX | PR | 24-AUG-2001; 2001US-31474P. |
| XX | PR | 22-JAN-2002; 2002US-35122P. |
| XX | PR | 29-MAR-2002; 2002US-368799P. |
| XX | PA | (INCY-) INVITE GENOMICS INC. |
| XX | PR | Tang YT, Yue H, Azimzai Y, Baughn MR, Burford N, Reddy R, Walia NK, Das D, Nguyen DB, Yao MG, Arvizu GS, Lu Y, Gandhi AR, Griffin JA, Elliott VS, Ramkumar J, Lal PG, Lu DAM, Lee EA, Lee SY, Yue H, Yang J, Tribouley CM, Kable AE, Svarnakar A; WPI; 2003-12079/11; P-PSDB; AAB34440. |
| XX | PT | The present invention relates to novel human lipid-associated molecules (LIPAM) and polyribonucleotides encoding such proteins. Sequences of the invention are useful for treating diseases or conditions associated with decreased expression of functional LIPAM. The antagonist is useful for treating a disease or condition associated with the overexpression of |
| XX | CC | dominance; priion disease; Creutzfeld-Jakob disease; anaemia; Alzheimer's disease; asthma; adenocarcinoma; lymphoma; melanoma; myeloma; sarcoma; gene therapy; protein replacement therapy; gene; ss. |
| XX | OS | Homo sapiens. |

functional LIPAM. They are useful for diagnosing, treating or preventing cardiovascular disorders (e.g. arteriovenous fistulae, atherosclerosis, heart failure, thrombophlebitis, angina pectoris, ischaemic heart disease or rheumatic heart disease), gastrointestinal disorders (e.g. peptic oesophagitis, nausea, peptic ulcer or Crohn's disease), lipid metabolism disorders (e.g. fatty liver, Fabry's disease, Gaucher's disease, diabetes mellitus, hyperlipidaemia, hypercholesterolaemia), autoimmune disorders or inflammatory disorders (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's disease), neurological disorders (e.g. stroke, epilepsy, dementia, Alzheimer's disease, or prion diseases such as kuru or Creutzfeldt-Jakob disease) or cancers (e.g. adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma or sarcoma). They are also used in gene therapy and protein replacement therapy. The present sequence is human LIPAM-1 cDNA.

| | | | | | | | |
|----|------|--|------|----|------|--|------|
| QY | 781 | TGGAAACAGCCTCCCTGGCCCTCCGAGGTAACATGAGCAGGAGTCCTCCTCACCGTGTTC | 840 | QY | 1861 | GAAGATTTCATCTGGTTGCAACCCTCTTGAAACAGTGCACATCCAAAGACCTG | 1920 |
| Db | 789 | TGGAAACAGCCTCCCTGGCCCTCCGAGGTAACATGAGCAGGAGTCCTCCTCACCGTGTTC | 848 | Db | 1869 | GAGATTTCATCTGGTTGCAACCCTCTTGAAACAGTGCACATCCAAAGACCTG | 1928 |
| QY | 841 | CAGCCTTCTTCTTATGAGACCCACCCCATCTTCACTGAGCAGGAGTCCTCCTCACCGTGTTC | 900 | QY | 1921 | GAAGATTGCCTGACAACCTCTTCTGGCTCTGACTGAGTCCTCCTCACCTGAGCAAGCT | 1980 |
| Db | 849 | CAGCCTTCTTCTTATGAGACCCACCCCATCTTCACTGAGCAGGAGTCCTCCTCACCGTGTTC | 908 | QY | 1929 | GAGGATTCAGTGCCTGACAACCTCTTCTGGCTCTGACTGAGTCCTCCTCACCTGAGCAAGCT | 1988 |
| QY | 901 | TCTAACAGCTGGCTGGCATCTTCTGGATAAGATGATGAGCTGGAGCAGGAGAAGAT | 960 | QY | 1981 | CACTCCGAGCAGCCAGTGTCTCTGAGAACATATGCTGGACCTGTTGGCAGAGACG | 2040 |
| Db | 909 | TCTAACAGCTGGCTGGCATCTTCTGGATAAGATGATGAGCTGGAGCAGGAGAAGAT | 968 | Db | 1989 | CACTCCGAGCAGCCAGTGTCTCTGAGAACATATGCTGGAGCAGAGACG | 2048 |
| QY | 1021 | CTCTTACAGTACAGAACAGCAACTACTTGACAGACTGCAAAACCCAAACAGCT | 1080 | QY | 2041 | ACTCGTCTTAAGTTGAACACAGATCATCACATGCCGACCCCTGGATGGTTC | 2100 |
| Db | 961 | CTCTTACAGTACAGAACAGCAACTACTTGACAGACTGCAAAACCCAAACAGCT | 1020 | Db | 2049 | ACTCGTCTTAAGTTGAACACAGATCATCACATGCCGACCCCTGGATGGTTC | 2108 |
| Db | 969 | GAGGCCATTGAGTGTAAACACGGGAGGCAATGAGTGTCTCTCAGAGAAGAT | 1028 | QY | 2101 | CTGAGGACTCTACAAGAACAGCAGTGCAGGAGTCATGGACCTGCTGCAAGA | 2160 |
| QY | 1081 | GGCTTAAGAGGAGGAGGAAATCAGATGCTCTGACAAGAACCCCTGGATGGTTC | 1140 | Db | 2109 | CTGAGGACTCTACAAGAACAGTGCAGGAGTCATGGACCTGCTGCAAGA | 2168 |
| Db | 1089 | GGCTTAAGAGGAGGAAATCAGATGCTCTGACAAGAACCCCTGGATGGTTC | 1148 | QY | 2221 | GTGCTCTGCTCTGGGGATCTCTGACCCCTGGATGGACCTGCTGCTGCAAGA | 2280 |
| QY | 1141 | ACCTCTAGTTCATAGCTGAGGCCGCTGACATCAACGTAATGAGCCTGCTGACTCT | 1200 | Db | 2229 | GTGCTCTGCTCTGGGGATCTCTGACCCCTGGATGGACCTGCTGCTGCAAGA | 2288 |
| Db | 1149 | ACCTCTAGTTCATAGCTGAGGCCGCTGACATCAACGTAATGAGCCTGCTGACTCT | 1208 | QY | 2281 | CTCCCCGAGTGTACACACAGTACGGAGCTGCTGACCCCTACCTCAGTGCAG | 2340 |
| QY | 1201 | CTCACGGCAGGCAATGGGGCCGGTCCACACTGGACAGTCAGTGCACCGTACCTG | 1260 | Db | 2289 | CTCCCCGAGTGTACACACAGTACGGAGCTGCTGACCCCTACCTCAGTGCAG | 2348 |
| Db | 1209 | CTCACGGCAGGCAATGGGGCCGGTCCACACTGGACAGTCAGTGCACCGTACCTG | 1268 | QY | 2341 | CTGGAGAATGTCRACACCTTACCTAATTCCTCGGAGTTCACAGAACCTCACAGC | 2400 |
| QY | 1261 | TACCGAGGCTGCTCTGGGGCGAGTGTACAGACATGGCACCGTACCTGACCTG | 1320 | Db | 2349 | CTGGAGAATGTCRACACCTTACCTAATTCCTCGGAGTTCACAGAACCTCACAGC | 2408 |
| Db | 1329 | GCGACATCTCCCGGAATTCAACCTCTCTGAGGCTTCTGTTGGAGCTGGAAA | 1388 | QY | 2401 | TACCCGTTGGCGCGGGTGTACGCAATCACAGAACCTCACAGTGCAGGAGGACGGTCC | 2460 |
| QY | 1381 | GAACACAGCTCTTATGCCCTTAAACAGGCTGTCGGAGGAGGCCGAGCTAGGATCTA | 1440 | Db | 2409 | TAGCCGTTGGGAGCAGGGTGTACGCAATCACAGAACCTCACAGTGCAGGAGGACGGTCC | 2468 |
| Db | 1389 | TACCGAGGCTCTTAAACAGGCTGTCGGAGCTGTAGAGATGACAGGAGTACCTTCG | 1500 | QY | 2451 | GGGCAAAGGCTGAGGATCTTATGAGGCAAGTCCAAACCTGTGAGAAGTGAAGAT | 2520 |
| QY | 1441 | CTCTCCAGGAGGAGCTGTCGGAGCTGTAGAGATGACAGGAGTACCTTCG | 1508 | Db | 2459 | GGGCAAAGGCTGAGGATCTTATGAGGCAAGTCCAAACCTGTGAGAAGTGAAGAT | 2528 |
| Db | 1449 | CTCTCCAGGAGGAGCTGTCGGAGCTGTAGAGATGACAGGAGTACCTTCG | 1508 | QY | 2521 | GATCATAGTAAATTTCATGAGAAGCTGGAGGTCTGAGAGGTCAGTGCAGGAGG | 2580 |
| QY | 1501 | GAAGACTGGAAATAACCTCTTATAGGGCAATGACCTCTGTGATTCGTCAAT | 1560 | Db | 2529 | GATCATAGTAAATTTCATGAGAAGCTGGAGGTCTGAGAGGTCAGTGCAGGAGG | 2588 |
| Db | 1509 | GAAGACTGGAAATAACCTCTTATAGGGCAATGACCTCTGTGATTCGTCAAT | 1568 | QY | 2551 | GATTATGCTGACTCTACATGAGAAGCTGGAGGTCTGAGAGGTCAGTGCAGGAGG | 2640 |
| QY | 1561 | GATCTGGTCACTATTCCTCCCAGAACTTCACAGACAACTATGAAAGGCCCTGAGACATC | 1620 | Db | 2559 | GATTTATGCTGACTCTACATGAGAAGCTGGAGGTCTGAGAGGTCAGTGCAGGAGG | 2648 |
| Db | 1569 | GATCTGGTCACTATTCCTCCCAGAACTTCACAGACAACTATGAAAGGCCCTGAGACATC | 1628 | QY | 2641 | CTCGCAATGCTTGGAGGCTCTGAGAGGTCAGTGCAGGAGGTCAGTGCAGGAGG | 2700 |
| QY | 1621 | CTCTCATGCTGAGGTCTCTGGGCTTGTGAGCTGGAGCTGGCTTGTAGTGTCACT | 1680 | Db | 2649 | GATTATGCTGACTCTACATGAGAAGCTGGAGGTCTGAGAGGTCAGTGCAGGAGG | 2708 |
| Db | 1629 | CTCTCATGCTGAGGTCTCTGGGCTTGTGAGCTGGAGCTGGCTTGTAGTGTCACT | 1688 | QY | 2701 | GACTCTCTGAGGAGCTTGTGAGCTGGAGCTGGCTTGTAGTGTCACT | 2760 |
| QY | 1741 | TGTCCTCTGCTCTGAGGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG | 1800 | Db | 2709 | GACTCTCTGAGGAGCTTGTGAGCTGGAGCTGGCTTGTAGTGTCACT | 2768 |
| Db | 1681 | CTGAGGGAGCTGTACCGAGGAAAGTCTACTGCCCCAGGAGTACTCTGGCTCTG | 1740 | QY | 2761 | CTCGAGGGCCGCTTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGG | 2880 |
| QY | 1749 | CTGAGGGAGCTGTACCGAGGAAAGTCTACTGCCCCAGGAGTACTCTGGCTCTG | 1748 | Db | 2769 | CTGAGGGAGCTGTACCGAGGAAAGTCTACTGCCCCAGGAGTACTCTGGCTCTG | 2888 |
| QY | 1801 | ACAGAGAAGTCTGAGGAGACCACACTGATGAGAGTGGGGATATGACACAGG | 1860 | QY | 2821 | CTGAGGGAGCTGTACCGAGGAAAGTCTACTGCCCCAGGAGTACTCTGGCTCTG | 2940 |
| Db | 1809 | ACAGAGAAGTCTGAGGAGACCACACTGATGAGAGTGGGGATATGACACAGG | 1868 | Db | 2829 | CTGAGGGAGCTGTACCGAGGAAAGTCTACTGCCCCAGGAGTACTCTGGCTCTG | 2948 |
| QY | 2941 | ATCCAGCTCCCTCTGGGGATACGTCCTCTTGGCCAGACTG | 3000 | | | | |

Db 2949 ATCCAGCTCCCTGTCGGCGATGGGCTCCAGATACTGCTCTTGGCCAGTC 3008
 Qy 3001 ATCCACCAAAATCAGAATTCACTCAGCTGGAGACGCCCTTGACCAATGCT 3060
 Db 3009 ATCACCACAAATCAGAAATTCACTCCAGCTGGAGACGCCCTTGACCAATGCT 3068
 Qy 3061 GAAACACTTGAGCAAACAGAGACCTGGAGCTGAGAGAGATGCCATCACCTG 3120
 Db 3069 GAAACACTTGAGCAAACAGAGACCTGGAGCTGAGAGAGATGCCATCACCTG 3128
 Qy 3121 CCACACTCGAAATGAGCCCTCTCTGAGAACCCCTGGAGATAGRACTACAGA 3180
 Db 3129 CCACACTCGAAATGAGCCCTCTGAGAACCCCTGGAGATAGRACTACAGA 3188
 Qy 3181 AGGCAGCCATGAGAAGCTGGCGAGCTACTCTGTGTAGTAACTACAGA 3240
 Db 3189 AGGCAGCCATGAGAAGCTGGCGAGCTACTCTGTGTAGTAACTACAGA 3248
 Qy 3241 AGTGTTCACCTCTGTGCCACCGAGCTCGACAGCAGACATCAAGTGTGCC 3300
 Db 3249 AGTGTTCACCTCTGTGCCACCGAGCTCGACAGCAGACATCAAGTGTGCC 3308
 Qy 3301 GGTGACTCTCTGACTACAGCACTGGGACTCTGGGACTCCACAAACACTC 3360
 Db 3309 TCTTGAGGGGACTCTGAGGAGCTGGGAGCTGGGAGCTGGGACTACACCA 3368
 Qy 3361 TCTTGAGGGGACTCTGAGGAGCTGGGAGCTGGGACTACACCA 3420
 Db 3369 TCTTGAGGGGACTCTGAGGAGCTGGGAGCTGGGACTACACCA 3428
 Qy 3421 CTGGCCAAATCTGAAGAGTCACCCCTACCTCCCTGGCTCTTACCGCACCTG 3480
 Db 3429 CTGGCCAAATCTGAAGAGTCACCCCTACCTCCCTGGCTCTTACCGCACCTG 3488
 Qy 3481 GGGGGGACAGCAGGACTAAATGGCAGGGAGGGGAGCTGGACTAGGCC 3540
 Db 3489 GAGGGGACAGGAGCTAAATGGCAGGGAGGGGAGCTGGACTAGGCC 3548
 Qy 3541 CGGGCCTGGGACTCTGTGAGGGAGTAACAGCCCGAGCATCACCTGGAGA 3600
 Db 3549 CGGGCCTGGGACTCTGTGAGGGAGTAACAGCCCGAGCATCACCTGGAGA 3608
 Qy 3601 TGGAGCTGGTCAACTCTCATGGGTCAAGACTGTGTCATACTGTGAGAATCG 3660
 Db 3609 TGGAGCTGGTCAACTCTCATGGGTCAAGACTGTGTCATACTGTGAGAATCG 3668
 Qy 3661 GGGCCCACCTGGCCACGGATAATGTTICAGCICATCCACAGGCCCTGGACATCCCT 3720
 Db 3669 GGGCCCACCTGGCCACGGATAATGTTICAGCICATCCACAGGCCCTGGACATCCCT 3728
 Qy 3721 GGGGAGCTCCAGGGTTCTCAAGTGGGGAGTCAGGAGCTGGCTGGCTCTAC 3780
 Db 3729 GGGGAGCTCCAGGGTTCTCAAGTGGGGAGTCAGGAGCTGGCTGGCTCTAC 3788
 Qy 3781 CGGGGCAAGGGGAATGTCATCTGGAGCTGGAGACACTGCACTTGCTCAGA 3840
 Db 3789 CAGGGCAAGGGGAATGTCATCTGGAGCTGGAGCTGGAGACACTGCACTTGCTCAGA 3848
 Qy 3841 CHTCGCAAASTCTCTGGAGAGCAACTGAGAAGTCACTGCACTTGCTCAGA 3900
 Db 3849 CACTCGCAAGGCTCTGAGAGCAACTGAGAAGTCACTGCACTTGCTCAGA 3908
 Qy 3901 GCGATCTCGCAAGGCTCTGAGAGCAACTGAGAAGTCACTGCACTTGCTCAGA 3960
 Db 3909 GCGATCTCGCAAGGCTCTGAGAGCAACTGAGAAGTCACTGCACTTGCTCAGA 3968
 Qy 3961 GTGGCAGCTTCTCCAAACCACTACCCACTGAGGAGAGGAGCTGACCTC 4020
 Db 3969 GTGCAGCTTCTCCAAACCACTACCCACTGAGGAGAGGAGCTGACCTC 4028
 Qy 4021 ACCTCTCTCGAGGACTGTTCACCTCTGAGCAGGCGCATCGGAGATGSCATC 4080

Db 4029 ACCTCTCTCCGAGGACTGTTCACCTCTGAGCCGGCATGGGAGATGGCCTC 4088
 Qy 4088 GACTCTGGAAACACATGCTGGCTGAGCTGAGCTGAGCTTCAACTTCAC 4140
 Db 4089 GCACTCTGGAAACACATGCTGGCTGAGCTGAGCTGAGCTTCAACTTCAC 4148
 Qy 4141 CACAGCCGAGCAACTCACTGGCCCTCTCTGGAGAGCCCTTACCTTACCCCTGG 4200
 Db 4149 CACAGCCGAGCAACTCACTGGCCCTCTCTGGAGAGCCCTTACCTTACCCCTGG 4208
 Qy 4201 AACAGCCGAGTGTCCAGACAGGGCTGAGAGGCCCTGAGCTGAGCTGAGCTC 4260
 Db 4209 AACAGCCGAGTGTCCAGACAGGGCTGAGAGGCCCTGAGCTGAGCTGAGCTC 4268
 Qy 4261 CGATGCGAGGGAGCTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTC 4320
 Db 4269 CGATGCGAGGGAGCTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTC 4328

Qy 4321 AGGAGGGTGGCGGAGGAGATCTCCAAATGAGCTGGCTGGCACTGTGCCCTCTAG 4377
 Db 4329 AGGAGGGTGGCGGAGGAGATCTCCAAATGAGCTGGCTGGCACTGTGCCCTCTAG 4385

RESULT 4
 ADD52634
 ID ADD52634 standard; cDNA; 4424 BP.
 XX
 DE Human lipid-associated molecule (LIPAM)-9 cDNA.
 XX
 KW Human; lipid-associated molecule; LIPAM; cardiovascular disorder; stroke;
 KW arteriovenous fistula; atherosclerosis; hypertension; Raynaud's disease;
 KW aneurysm; congestive heart failure; thromboembolism; angina pectoris;
 KW ischaemic heart disease; rheumatic heart disease; peptic oesophagitis;
 KW gastrointestinal disorder; lipid metabolism disorder; Crohn's disease;
 KW nausea; peptic ulcer; fatty liver; Fabry's disease; Gaucher's disease;
 KW diabetes mellitus; hyperlipidaemia; hypercholesterolaemia; epilepsy;
 KW autoimmune disorder; inflammatory disorder; anaemia; Alzheimer's disease; asthma;
 KW dementia; prion disease; Creutzfeldt-Jakob disease; leukaemia; cancer;
 KW adenocarcinoma; lymphoma; melanoma; myeloma; sarcoma; gene therapy;
 KW protein replacement therapy; gene; ss.
 XX
 OS Homo sapiens.

XX
 FT Key Location/Qualifiers
 FT CDS 9..2918
 FT /product= "Human LIPAM-9 protein"
 FT sig_peptide 9..65
 FT mat_peptide 66..2915
 FT /*tag= a
 FT /*tag= c
 FT /product= "Human mature LIPAM-9 protein"
 FT sig_peptide 9..71
 FT /*tag= d
 FT mat_peptide 72..2915
 FT /*tag= e
 FT /product= "Human mature LIPAM-9 protein"
 FT sig_peptide 9..74
 FT /*tag= f
 FT mat_peptide 75..2915
 FT /*tag= g
 FT /product= "Human mature LIPAM-9 protein"
 FT sig_peptide 9..83
 FT /*tag= h
 FT /*tag= i
 FT mat_peptide 75..2915
 FT /*tag= j
 FT /product= "Human mature LIPAM-9 protein"

PN WO20029498-B-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 17-MAY-2002; 2002WO-US15688.
 PR 18-MAY-2001; 2001US-292242P.
 PR 25-MAY-2001; 2001US-293726P.
 PR 01-JUN-2001; 2001US-295346P.
 PR 06-JUL-2001; 2001US-303404P.
 PR 24-AUG-2001; 2001US-314754P.
 PR 22-JAN-2002; 2002US-351262P.
 PR 29-MAR-2002; 2002US-368799P.
 PA
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Tang YT, Yue H, Azimzai Y, Baughn MR, Burford N, Reddy R;
 PI Walia NK, Das D, Nguyen DB, Yao MG, Arvizu CS, Lu Y, Gandhi AR;
 PI Griffin JA, Elliott VS, Ramkumar J, Lal PG, Lu DAM, Lee EA;
 PI Lee SY, Yue H, Yang J, Tribouley CM, Kable AE, Swarnakar A;
 XX
 DR WPI; 2003-120797/11.
 DR P-PSDB; AAE34448.
 XX
 PT New human lipid-associated molecule (LIPAM) proteins and
 PT polynucleotides, useful for diagnosing, treating or preventing
 PT cardiovascular disorders (e.g. aneurysms), neurological disorders (e.g.
 XX
 PS Claim 74; Page 168-169; 171pp; English.

The present invention relates to novel human lipid-associated molecules
 CC (LIPAM) and polynucleotides encoding such proteins. Sequences of the
 CC invention are useful for treating diseases or conditions associated with
 CC decreased expression of functional LIPAM. The antagonist is useful for
 CC treating a disease or condition associated with the overexpression of
 CC functional LIPAM. They are useful for diagnosing, treating or preventing,
 CC cardiovascular disorders (e.g. arteriovenous fistula, atherosclerosis,
 CC hypertension, Raynaud's disease, varicose veins, congestive
 CC heart failure, thrombophlebitis, angina pectoris, ischaemic heart disease
 CC or rheumatic heart disease), gastrointestinal disorders (e.g. peptic
 CC oesophagitis, nausea, peptic ulcer or Crohn's disease), lipid metabolism
 CC disorders (e.g. fatty liver, Fabry's disease, Gaucher's disease, diabetes
 CC mellitus, hyperlipidaemia, hypercholesterolaemia), autoimmune disorders
 CC or inflammatory disorders (e.g. acquired immunodeficiency syndrome,
 CC anaemia, asthma or Cohn's disease), neurological disorders (e.g. stroke,
 CC epilepsy, dementia, Alzheimer's disease, or prion diseases such as kuru
 CC or Creutzfeldt-Jakob disease) or cancers (e.g. adenocarcinoma, leukaemia,
 CC lymphoma, melanoma, myeloma or sarcoma). They are also used in gene
 CC therapy and protein replacement therapy. The present sequence is human
 XX
 Sequence 4424 BP; 1119 A; 1231 C; 1143 G; 931 T; 0 other;

Query Match 91.0%; Score 3893.4; DB 25; Length 4424;
 Best Local Similarity 95.6%; Pred. No. 0; Mismatches 6; Indels 187; Gaps 1;
 Matches 4184; Conservative 0;

QY 1 ATGGGGCTCGGCCAGGATTTCCTCCGTGAGCTGCTCTGCTCTGGGCCAGGCC
 QY 9 ATGGGGCTCGGCCAGGATTTCCTCCGTGAGCTGCTCTGCTCTGGGCCAGGCC 60
 Db 61 CTCAGATCCATACCTCTCTAGAAGAGCTAATGGGAAGCTGAGCTGCTCTGGGCCAGGCC 120
 Db 69 CTCAGATCCATACCTCTCTAGAAGAGCTAATGGGAAGCTGAGCTGCTCTGGGCCAGGCC 128
 QY 121 CTGAGAAATTCCTCCATGCAACCAAATAATTGGAGTAATGCTCTCAA 180
 Db 123 CTGAGAAATTCCTCCATGCAACCAAATAATTGGAGTAATGCTCTCAA 188
 QY 181 TAGTCACTCTGAGGCTCTGATATAATTGGAGCTGCAATCTGAA 240
 Db 189 TAGTCACTCTGAGGCTCTGCAAGCTCTGATATAATTGGAGCTGCAATCTGAA 248

OY 241 ATTCTCCAGACCCAGGGACGGGGGATCTGGAGAACAGACTGGACTGAAGGCCAG 300
 OY 249 ATTCCCTCAGACCCAGGGACGGGGATCTGGAGAACAGACTGGACTGAAGGCCAG 308
 Db 301 CAGGTGTCAGCTGGAGTAGATCACAGTCAGCTTCAGACATCACAGATATTTCAGTCCT 360
 OY 309 CAGGTGTCAGCTGGAGTAGATCACAGTCAGCTTCAGACATCACAGATATTTCAGTCCT 368
 Db 361 GTTCCAAATGCCCTGTCACACTGGAGAGTCATACCCAGATGGCTCTGAGAC 420
 OY 369 GTTCCAAATGCCCTGTCACACTGGAGAGTCATACCCAGATGGCTCTGAGAC 428
 Db 421 TTCTGGATTCAAGCTCAGAATGGCTTGCGGGGGGGGGGGATGAGCTGCAACTGACTT 480
 OY 429 TTCTGGATTCAAGCTCAGAATGGCTTGCGGGGGGGGGATGAGCTGCAACTGACTT 488
 Db 481 CTTTGTGACTGGAGTCATCATGTTCTTCAGTATGTCAGCTTACGCCAGTGTACCTGTC 540
 OY 489 CTTTGTGACTGGAGTCATCATGTTCTTCAGTATGTCAGCTTACGCCAGTGTACCTGTC 548
 Db 541 CCCCTCTGTCACAGAAATGGCTTGCGGGGGGGGGGGATGAGCTGCAACTGACTT 600
 OY 549 CCTCTCTGTCACAGAAATGGCTTGCGGGGGGGGGGGATGAGCTGCAACTGACTT 608
 Db 601 GACTACTTGAGGAGGTCCCAGAGCATTTGAAACCTCTGGACTCTCTGAGTT 660
 OY 609 GACTACCTGCCAGGAGGTCCCAGAGCATTTGAAACCTCTGGACTCTCTGAGTT 668
 Ps 661 GCAGAGGTCTCTGTCAGTATACGGCTACTCTGCTCAGCCCTGCAAGACSCCTTAT 720
 OY 669 GCAGAGGTCTCTGTCAGTATACGGCTACTCTGCTCAGCCCTGCAAGACSCCTTAT 728
 Db 721 TGCTCAGGGAGACCCCGTGCCAGAGGTCTCTGAGCTTCTGAGGAGTCCTCACCCTGTT 780
 OY 729 TCTCTGAGGAGACCCCGTGCCAGGAGTCCTCTGAGGAGTCCTCACCCTGTT 788
 Db 781 TGGACACACCTCTGGCTCAGGACTACTGAGGAGTCCTCTGAGGAGTCCTCACCCTGTT 840
 OY 789 TGGACACCTCTGGCTCAGGACTACTGAGGAGTCCTCACCCTGTT 848
 Db 841 CACGCTTCTCTPATGAGACCCATCTCTACACTGGAGACCCGACTCCAGAT 900
 OY 849 CACGCTTCTCTPATGAGACCCATCTCTACACTGGAGACCCGACTCCAGAT 908
 Db 901 TCTACCAAGCTGCCTGCATCTCTGAGATGGATGGAGGCCACAGGAGAGAT 960
 OY 909 TCTACCAAGCTGCCTGCATCTCTGAGATGGATGGAGGCCACAGGAGAGAT 968
 Db 961 GACCCATTGAGCTAAACCGGAGCCAATGAGTGTCCCTCTCAGGAGGCCCT 1020
 OY 969 GACCCATTGAGCTAAACCGGAGCCAATGAGTGTCCCTCTCAGGAGGCCCT 1028
 Db 1021 CTGTCCTACTACAGAAGACGACTACCTGACAGACTGAGCTGAGAACCCAGACAGCTT 1080
 OY 1029 CTGTCCTACTACAGAAGACGACTACCTGACAGACTGAGCTGAGAACCCAGACAGCTT 1088
 Db 1031 GAGGTGAGGAGGAGGAGGAATCAGTGTCTGAGAACGCTCCGATACGGTCCC 1140
 OY 1039 GAGGTGAGGAGGAGGAGGAATCAGTGTCTGAGAACGCTCCGATACGGTCCC 1148
 Db 1141 ACCTCACTGTCAGGCTGAGCGGGCTGACATCACAGTAATGGAGCCCTGGTACT 1200
 OY 1149 ACTTCAGATCTCATGGCTCAAGCGGTGACATCACATTATGGAGCCCTGGTACT 1208
 Db 1201 CTGAGAAATTCCTCCATGCAACCAAATAATTGGAGTAATGCTCTCAA 1260
 OY 1209 CTACGGAGGCAATGGGCCGGTCCACCTGGAGAACGCTGAGCTCTGAGCT 1268
 Db 1211 TACGGAGGCAATGGGCCGGTCCACCTGGAGAACGCTGAGCTCTGAGCT 1320
 OY 1251 TACGGAGGCAATGGGCCGGTCCACCTGGAGAACGCTGAGCTCTGAGCT 1328
 Db 1259 TACGGAGGCAATGGGCCGGTCCACCTGGAGAACGCTGAGCTCTGAGCT 1328

| | | | | |
|----|------|---|------|----|
| QY | 1321 | GCGAACATCCTCCGGAAATCAACCCCTCCGAACTCAGGGGTCTCTGCGCATGGAAA | 1388 | Db |
| Db | 1329 | GCGAACATCCTCCGGAAATCAACCCCTCCGAACTCAGGGGTCTCTGCGCATGGAAA | 1388 | Db |
| QY | 1381 | GAAACAGCTCTAATGCCCTCTTAAACCGGCTGGGAGGAGCCGAGCTGAGGATCTA | 1440 | Db |
| Db | 1389 | GAAACAGCTCTAATGCCCTCTTAAACCGGCTGGGAGGAGCCGAGCTGAGGATCTA | 1440 | Db |
| QY | 1441 | CCTGTCAGGCCAGGAGGCTGGGACCTGATGAGAATGACCGGAGATACTTCAG | 1500 | Qy |
| Db | 1449 | CCTGTCAGGCCAGGAGGCTGGGACCTGATGAGAATGACCGGAGATACTTCAG | 1508 | Qy |
| QY | 1501 | GAAAGCTGGAGATAATAACCTGTTATAGGCGGAATGACCTCTGATGCAAT | 1560 | Qy |
| Db | 1509 | GAAAGCTGGAGATAATAACCTGTTATAGGCGGAATGACCTCTGATGCAAT | 1568 | Qy |
| QY | 1561 | GATCTGGCAACTATCTCCCGAGAACTCACAGAACATGACCTCTGATGCAAT | 1620 | Qy |
| Db | 1569 | GATCTGGCAACTATCTCCCGAGAACTCACAGAACATGACCTCTGATGCAAT | 1628 | Qy |
| QY | 1621 | CTTCATGCTGAGGTTCTCGGCATTGGAACCTGGTGACGGCTCTGAGATGTCAC | 1680 | Db |
| Db | 1629 | CTTCATGCTGAGGTTCTCGGCATTGGAACCTGGTGACGGCTCTGAGATGTCAC | 1688 | Qy |
| QY | 1681 | CTGAGGGACTGTACCGAGAAAAAGTCACTGCCCCAGGATGATCTCAGCTCTG | 1740 | Db |
| Db | 1689 | CTGAGGGACTGTACCGAGAAAAAGTCACTGCCCCAGGATGATCTCAGCTCTG | 1748 | Qy |
| QY | 1741 | TGRCCTGCTGCTGAGTTGATGTAACCTCACAGAACTTGTACCTCATGAAATC | 1800 | Db |
| Db | 1749 | TGRCCTGCTGCTGAGTTGATGTAACCTCACAGAACTTGTACCTCATGAAATC | 1808 | Qy |
| QY | 1801 | AACAGAACTTCAGGAGAACCCACCACTGATGAGGTGGGATATGACAAG | 1860 | Db |
| Db | 1809 | AACAGAACTTCAGGAGAACCCACCACTGATGAGGTGGGATATGACAAG | 1868 | Qy |
| QY | 1861 | GAAGATTACTGCTGCTGAGCCCGTTTGAACAGCTGGACATGCCAAGACCTCG | 1920 | Db |
| Db | 1869 | GAAGATTACTGCTGCTGAGCCCGTTTGAACAGCTGGACATGCCAAGACCTCG | 1928 | Qy |
| QY | 1921 | GAAGATTACTGCTGCTGAGCCCGTTTGAACAGCTGGACATGCCAAGACCTCG | 1980 | Db |
| Db | 1929 | GAAGATTACTGCTGAGCCCGTTTGAACAGCTGGACATGCCAAGACCTCG | 1988 | Qy |
| QY | 1981 | CACTCCGAGCAGCCAGCTCTGGACATATGCTGGAGCTGTTGCCAGAGAG | 2040 | Db |
| Db | 1989 | CACTCCGAGCAGCCAGCTCTGGACATATGCTGGAGCTGTTGCCAGAGAG | 2048 | Qy |
| QY | 2041 | ACTCTCATAGTTGAGAACAGAACATCAATATCACATGTCGACCTTGCCAG | 2100 | Db |
| Db | 2049 | ACTCTCATAGTTGAGAACAGAACATCAATATCACATGTCGACCTTGCCAG | 2108 | Qy |
| QY | 2101 | CTGAGGACTACAGAACAGCATGGACCTGGCTCCATCGGACAGCGCTT | 2160 | Db |
| Db | 2109 | CTGAGGACTACAGAACAGCATGGACCTGGCTCCATCGGACAGCGCTT | 2168 | Qy |
| QY | 2161 | GCCCCCTCGCCTGCACCTACCTCAGTCATGCCCTGAGACCTGACACATCAAGT | 2220 | Db |
| Db | 2169 | GCCCCCTCGCCTGCACCTACCTCAGTCATGCCCTGAGACCTGACACATCAAGT | 2228 | Qy |
| QY | 2221 | GTCGCTGCTGCTGAGGATCTCTAACCGTGGCATGGATTCGCTCAAACCGAC | 2280 | Db |
| Db | 2229 | GTCGCTGCTGCTGAGGATCTCTAACCGTGGCATGGATTCGCTCAAACCGAC | 2288 | Qy |
| QY | 2281 | CTCCCGAGTCAACACAGCATGGACCTGGACATGGATTCGCTCAAACCGAC | 2340 | Db |
| Db | 2289 | CTCCCGAGTCAACACAGCATGGACCTGGACATGGATTCGCTCAAACCGAC | 2348 | Qy |
| QY | 2341 | CTGGGAATGTGACCACTTACCTATCTGGAGTTAACGAAACCTCACAGC | 2400 | Db |
| Db | 2349 | CTGGGAATGTGACCACTTACCTATCTGGAGTTAACGAAACCTCACAGC | 2408 | Qy |
| QY | 2401 | TACGCCGCTGGCACGGGTATGCCATGACAGCATGGATTCCTCAATGAGCTTCC | 2460 | Db |
| QY | 2409 | TACGCCGCTGGCACGGGTATGCCATGACAGCATGGATTCCTCAATGAGCTTCC | 2468 | Db |
| QY | 2461 | GAGGAAAGGCTGAGGATCTTGTAGGCCAAGTCACACTCTGATGCGAGAGAT | 2520 | Db |
| Db | 2469 | GAGGAAAGGCTGAGGATCTTGTAGGCCAAGTCACACTCTGATGCGAGAGAT | 2528 | Db |
| QY | 2521 | GATCTAGTAAATTCATGAGAATGAGCTGAGGTCATCACAGTGATGGGGCAGC | 2580 | Qy |
| Db | 2529 | GATCTAGTAAATTCATGAGAATGAGCTGAGGTCATCACAGTGATGGGGCAGC | 2588 | Db |
| QY | 2581 | GATTATGCACTACTGACAGATGAGATCTGATCTGAGTCTGAGCAACTTGTGACCAT | 2640 | Qy |
| Db | 2589 | GATTATGCACTACTGACAGATGAGATCTGATCTGAGTCTGAGCAACTTGTGACCAT | 2648 | Db |
| QY | 2641 | CTCCGAATGCCCTGGAAGCTCTGCATAGAGGGTGGCCAGAGTCTGGTAACTCTGTG | 2700 | Qy |
| Db | 2649 | CTCCGAATGCCCTGGAAGCTCTGCATAGAGGGTGGCCAGAGTCTGGTAACTCTGTG | 2708 | Qy |
| QY | 2701 | GAECTCTGCAACCCACTATCACCGGAGGTGGTCTGGGAAACCCAGACAGTGCCA | 2760 | Db |
| Db | 2709 | GAECTCTGCAACCCACTATCACCGGAGGTGGTCTGGGAAACCCAGACAGTGCCA | 2768 | Qy |
| QY | 2761 | GTGCGAGGCCAGGCTTGTGTAACCGGTTCTGACCGCTGAGGAGTCCACCCCAG | 2820 | Db |
| Db | 2769 | GTGAGCAGGCCA | 2781 | Qy |
| QY | 2821 | CTAGCCAGGCTGGAGGCTTCAGCGAGCTACCGGAGCATGGCAGACTCCAAAGAG | 2880 | Db |
| Db | 2782 | - | 2781 | Qy |
| QY | 2881 | TCAGCCGCTATGACACCGAGGAGACTTCTCTGCTGAGCCCTTCTCCAGAAC | 2940 | Db |
| Db | 2782 | - | 2781 | Qy |
| QY | 2941 | ATCCAGCTCTGCTGGGATGGCTCCAGATACGCTCTCTCTTGGCCAGACTGC | 3000 | Db |
| Db | 2782 | - | 2781 | Qy |
| QY | 3001 | ATCCACCCAAATCAGAACATTCACTCCACTCCAGCTGGCCAGGCTTGGACCAATATGCT | 3060 | Db |
| Db | 2822 | ATCCACCCAAATCAGAACATTCACTCCAGCTGGCCAGGACCCCTTGGACCAATATGCT | 2881 | Qy |
| QY | 3061 | GAACACTCTGGAGAACAAAGAACCCCTGGACCTGAGGAGGATGCCATCACCTG | 3120 | Db |
| Db | 2882 | GAACACTCTGGAGAACAAAGAACCCCTGGACCTGAGGAGGATGCCATCACCTG | 2941 | Qy |
| QY | 3121 | CCCACTGAGAATGAGCCCTCTGAGAACGGCTGGAAATAGTAACGATGCCAT | 3180 | Db |
| Db | 2942 | CCCACTGAGAATGAGCCCTCTGAGAACGGCTGGAAATAGTAACGATGCCAT | 3001 | Qy |
| QY | 3181 | AAGCCAGCCATGAGAACCTGGGAGCTGACTCTGGTGTGAGAGTGGCTTCAAT | 3240 | Db |
| Db | 3002 | AAGCCAGCCATGAGAACCTGGGAGCTGACTCTGGTGTGAGAGTGGCTTCAAT | 3061 | Qy |
| QY | 3241 | AGTGTCTCAACTCTGCAACCACTCCGACACAGCACAGCATCAAGTGCGGCCCTG | 3300 | Db |
| Db | 3062 | AGTGTCTCAACTCTGCAACCACTCCGACACAGCACAGCATCAAGTGCGGCCCTG | 3121 | Qy |
| QY | 3301 | GGTAGCTCTGACTACAGGAGGTGGGAGCTGAGCAACACATCCAGTGAACCTACCA | 3360 | Db |
| Db | 3122 | GGTAGCTCTGACTACAGGAGGTGGGAGCTGAGCAACACATCCAGTGAACCTACCA | 3181 | Qy |
| QY | 3361 | TCTTGAGGAGACTCTGGAGCATGGGAGCTGGAGCTGAGCAACACATCCAGTGAACCTACCA | 3420 | Db |
| Db | 3182 | TCTTGAGGAGACTCTGGAGCATGGGAGCTGGAGCTGAGCAACACATCCAGTGAACCTACCA | 3241 | Qy |
| QY | 3421 | CTGCCAACATCTGAGAGGTCAACCTTACCTCTGGCTCTCTTACACCGCT | 3480 | Db |
| Db | 3242 | CTGCCAACATCTGAGAGGTCAACCTTACCTCTGGCTCTCTTACACCGCT | 3301 | Qy |
| QY | 3481 | GAGGGACAGGAGACTAATGTCAGGGAGGGAGGGAGGGAGCTGGAGACATCCAGCC | 3540 | Db |

| | | | | |
|-----------------|---|--|-------------------------------|--|
| Db | 3302 | GAGGGGAGAGCAGGACTAAATGTGCGAGGGCCAGAGCTAGGACATGCCAGCC | 3361 | OS Homo sapiens. |
| Oy | 3541 | CAGGCCTTGAGCCTTGAGAGTAAACAGGCGGAGCATCAACCTGAGAAGAC | 3600 | XK W020022757-A2. |
| Db | 3362 | CAGGCCTGGGACCTGTGAGAAACAGGCGGAGCATCAACCTGAGAAGAC | 3421 | XK PD 19-SEP-2002. |
| Oy | 3601 | TGGAGCTGTCAACTCTTCAATTGGGTCAGACACTGTGCAATTGTGAGAATCCAGCC | 3660 | PR XX 08-MAR-2002; 2002WO-US06908. |
| Db | 3422 | TGGAGCTGTCAACTCTTCAATTGGGTCAGACACTGTGCAATTGTGAGAATCCAGCC | 3481 | PR XX 08-MAR-2001; 2001US-274101P. |
| Oy | 3661 | GAGGCCCACTTGGCAGGAAATGTTCAGCACATCCACAGGCGCTGACATCTCT | 3720 | PR 08-MAR-2001; 2001US-274281P. |
| Db | 3482 | GAGGCCCACTTGGCAGGAAATGTTCAGCACATCCACAGGCGCTGACATCTCT | 3541 | PR 09-MAR-2001; 2001US-274849P. |
| Oy | 3721 | GAGGAGCTCCAAAGGCTTGTGACCTGGAGCTGAGCTGAGCTACAGCTGTTGAGAATCCAGCC | 3780 | PR 12-MAR-2001; 2001US-275235P. |
| Db | 3542 | GAGGAGCTCCAAAGGCTTGTGACCTGGAGCTGAGCTGAGCTACAGCTGTTGAGAATCCAGCC | 3601 | PR 13-MAR-2001; 2001US-275578P. |
| Oy | 3781 | CAGGCCCAAGGCGGAATGTGCGATGTCGGAGCTGAGAACAGGAAAGTGAACCTCCAGA | 3840 | PR 13-MAR-2001; 2001US-275601P. |
| Db | 3602 | CAGGCCCAAGGCGGAATGTGCGATGTCGGAGCTGAGAACAGGAAAGTGAACCTCCAGA | 3661 | PR 14-MAR-2001; 2001US-276000P. |
| Oy | 3841 | CACTCGCAAAGCTCCTGAGAACAGAAAGTGAACCTGGACCTCCAGA | 3900 | PR 16-MAR-2001; 2001US-276769P. |
| Db | 3662 | CACTCGCAAAGCTCCTGAGAACAGAAAGTGAACCTGGACCTCCAGA | 3721 | PR 19-MAR-2001; 2001US-276994P. |
| Oy | 3901 | GGCATCTCAGTTCTCTACTGGCACCATAACACAGCTGAGGCTTGGTTGAGAATCCAGA | 3960 | PR 20-MAR-2001; 2001US-277239P. |
| Db | 3782 | GTCGAGCCTTCTCTACTGGCACCATAACACAGCTGAGGCTTGGTTGAGAATCCAGA | 3841 | PR 20-MAR-2001; 2001US-277321P. |
| Oy | 4021 | ACCTCTTCTCGAGGACTGTTCACTCTAGACCGGGGATGCCAGATGCCATC | 4080 | PR 21-MAR-2001; 2001US-277791P. |
| Db | 3842 | ACCTCTTCTCGAGGACTGTTCACTCTAGACCGGGGATGCCAGATGCCATC | 3901 | PR 22-MAR-2001; 2001US-277833P. |
| Oy | 4081 | GCACTCTGAAACACATGCTGGAACCTGGCGCGCAGACTACCTCCACAACTTCA | 4140 | PR 23-MAR-2001; 2001US-278152P. |
| Db | 3902 | GCACTCTGAAACACATGCTGGAACCCACTGAAAGAGGGGACTGACCTCCACAACTTCA | 3961 | PR 26-MAR-2001; 2001US-278894P. |
| Oy | 4141 | CACAGCCGAACTCAAGTGCCTCTCTGAGASCCCTTAACCTCACCTGG | 4200 | PR 27-MAR-2001; 2001US-278999P. |
| Db | 3962 | CACAGCCGAACTCAAGTGCCTCTCTGAGASCCCTTAACCTCACCTGG | 4021 | PR 27-MAR-2001; 2001US-279342P. |
| Oy | 4201 | ACAGCCGAACTCAAGTGCCTCTCTGAGASCCCTTAACCTCACCTGG | 4260 | PR 28-MAR-2001; 2001US-279344P. |
| Db | 4022 | ACAGCCGAACTCAAGTGCCTCTCTGAGASCCCTTAACCTCACCTGG | 4081 | PR 30-MAR-2001; 2001US-277338P. |
| Oy | 4261 | CACAGCCGAACTCAAGTGCCTCTCTGAGASCCCTTAACCTCACCTGG | 4320 | PR 30-MAR-2001; 2001US-280233P. |
| Db | 4082 | CACAGCCGAACTCAAGTGCCTCTCTGAGASCCCTTAACCTCACCTGG | 4141 | PR 02-APR-2001; 2001US-280802P. |
| Oy | 4321 | AGGAGAGCTGGCGAGGAAAGTCTCAATAGGCTCGACTGTGCCCTAG | 4377 | PR 02-APR-2001; 2001US-280900P. |
| Db | 4142 | AGGAGAGCTGGCGAGGAAAGTCTCAATAGGCTCGACTGTGCCCTAG | 4198 | PR 04-APR-2001; 2001US-281194P. |
| RESULT 5 | | | | |
| ABX97050 | ID | ABX97050 Standard; cdna; 4425 bp. | 13-APR-2001; 2001US-283675P. | |
| XX | AC | ABX97050; | 16-MAY-2001; 2001US-2837185P. | |
| XX | DT | 20-MAY-2003 (first entry) | 02-MAY-2001; 2001US-288066P. | |
| XX | DE | Human NOV24c cDNA. | 03-MAY-2001; 2001US-288342P. | |
| KW | NOX; cytostatic; cardiotonic; antiarteriosclerotic; antiasthmatic; cancer; human; gene; ss. | 03-MAY-2001; 2001US-288528P. | | |
| KW | (CURA-) CURAGEN CORP. | 04-MAY-2001; 2001US-291190P. | | |
| XX | | 15-MAY-2001; 2001US-291999P. | | |
| | | 16-MAY-2001; 2001US-291999P. | | |
| | | 16-MAY-2001; 2001US-291240P. | | |
| | | 30-MAY-2001; 2001US-294485P. | | |
| | | 31-MAY-2001; 2001US-294889P. | | |
| | | 31-MAY-2001; 2001US-294999P. | | |
| | | 18-JUN-2001; 2001US-299027P. | | |
| | | 19-JUN-2001; 2001US-299303P. | | |
| | | 19-JUN-2001; 2001US-299310P. | | |
| | | 10-JUL-2001; 2001US-304354P. | | |
| | | 31-JUL-2001; 2001US-309198P. | | |
| | | 16-AUG-2001; 2001US-312903P. | | |
| | | 10-SEP-2001; 2001US-318462P. | | |
| | | 12-SEP-2001; 2001US-318770P. | | |
| | | 27-SEP-2001; 2001US-325430P. | | |
| | | 18-OCT-2001; 2001US-325681P. | | |
| | | 31-OCT-2001; 2001US-335301P. | | |
| | | 14-NOV-2001; 2001US-332172P. | | |
| | | 14-NOV-2001; 2001US-332271P. | | |
| | | 14-NOV-2001; 2001US-333184P. | | |
| | | 14-NOV-2001; 2001US-333272P. | | |
| | | 21-NOV-2001; 2001US-332094P. | | |
| | | 03-DEC-2001; 2001US-337426P. | | |
| | | 04-DEC-2001; 2001US-338092P. | | |
| | | 03-JAN-2002; 2002US-345705P. | | |
| | | 07-MAR-2002; 2002US-0092900. | | |

| | | | | | | | | | |
|----------------|--|-----------------|------------|-------------|-----------|-------|----|------|---|
| PI | Padigaru M, | Spytek KA, | Shenoy SC, | Taupier RJ, | Pena CEA, | Li L; | Db | 484 | CCCTTGCTACACAAATGGCTTGGGGGGGGGGGAGTTGACTGATGGGGTC |
| Zerniushen BD, | Gusev V, | Ji W, | Gorman L, | Miller CE, | Kekuda R, | | Db | 501 | GAATCTCAGGAGGAGCTCCCAGAGATTTGAAACCTGGTGACCTCTCTGAGTT |
| Patturajan M, | Ganguli E, | Vernet CAM, | Guo X, | Tchernev V; | | | Qy | 560 | |
| Fernandes ER, | Casman SJ, | Malyankar UM, | Garlich V, | Liu Y; | | | Db | 544 | GAATCTCAGGAGGAGCTCCCAGAGATTTGAAACCTGGTGACCTCTCTGAGTT |
| Anderson D, | Spaderna SK, | Catterton E, | Burgess C, | Leite M, | Zhong H; | | Db | 603 | |
| Alsobrook JP, | Lepley DM, | Rieger DK; | | | | | Qy | 661 | GGAGGGTCCTCGTAGATCAGGACTTGGTCAAGCTGCCAGAGGCCGTAAT |
| DR | WPI; | 2002-723332/78. | | | | | Db | 720 | |
| XX | P-PSDB; | ABU5083. | | | | | Qy | 664 | GCAGAGGTCTCTGTCAGATAGGACTGGCTAGCCAGAGGCCGTCATAAGGCC |
| PT | NOXV polypeptides and poly nucleotides, useful for preventing or | | | | | | Db | 666 | |
| PT | treating a disorder associated with aberrant NOX expression or | | | | | | Qy | 721 | TGCTCAGAGGAGACACCCCGTGGCAAGGGTATGGGATGGCTTACAGGAGCC |
| PT | activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or | | | | | | Db | 723 | |
| PT | bronchial asthma - | | | | | | Qy | 604 | TGCTCAGAGGAGACACCCCGTGGCAAGGGTATGGGATGGCTTACAGGAGCC |
| XX | Claim 13; Page 169-170; 1103pp; English. | | | | | | Db | 663 | |
| CC | This invention describes novel human NOXV polypeptides which have | | | | | | Qy | 724 | TGGAACAGCTCTCTGGCCCTCGAGCTAGAGCTGGTGGCTTCACGGTGGTT |
| CC | cystostatic, cardiotonic, antiarrhythmic, antidiastolic, antiasthmatic and | | | | | | Db | 780 | |
| CC | hypotensive activity. Pharmaceutical compositions comprising the NOXV | | | | | | Qy | 841 | AGGCTTCTTATAGAGACCCATCTACACTGGAGGACCCCCACTCGAGAT |
| CC | proteins or nucleic acid molecules or NOXV antibodies as useful for | | | | | | Db | 723 | |
| CC | preventing or treating a disorder associated with aberrant NOXV | | | | | | Qy | 781 | TGGAACAGCTCTCTGGCCCTCGAGCTAGAGCTGGTGGCTTCACGGTGGTT |
| CC | expression or activity e.g., cancer, hypertension, atherosclerosis, | | | | | | Db | 840 | |
| CC | cardiomyopathy or bronchial asthma. The products of the invention can | | | | | | Qy | 783 | |
| CC | be used for gene therapy or in a vaccine. ABX97008-ABX97105 are cDNA | | | | | | Db | 784 | CAGCTTCTCTATGAGACCCATCTACACTGGAGGACCCCCACTCGAGAT |
| CC | fragments amplified and isolated by the PCR primers and probes | | | | | | Qy | 831 | |
| CC | represented in ABX13460-ABX13462 and ABX97106-ABX9753. ABX97008-ABX97105 | | | | | | Db | 901 | TCTACACACCTGGCTGGATCTCTGGAATTAGGTATGGATGAGCTGAGCAGAGAT |
| CC | encode the NOXV proteins described in ABU65041-ABU65218. | | | | | | Qy | 960 | |
| XX | Sequence 4425 BP; 1089 A; 1242 C; 1173 G; 921 T; 0 other; | | | | | | Db | 832 | TCTACACACCTGGCTGGATCTCTGGAATTAGGTATGGATGAGCTGAGCAGAGAT |
| SQ | Query Match B9.6%; Score 3923; DB 24; Length 4425; | | | | | | Qy | 961 | GAGGATTGGATGTTAACACGGAGGGAGGGCATATGAGCTGAGCAGAGAT |
| | Best Local Similarity 95.4%; Pred No. 0; Mismatches 0; Indels 117; Gaps 9; | | | | | | Qy | 1020 | |
| | Matches 4181; Conservative 0; Mis matches 85; Indels 117; Gaps 9; | | | | | | Db | 892 | GAGCCATTGGTAAACACGGAGGGAGGGCATATGAGCTGAGCAGAGAT |
| Qy | 1 ATGGGGCTGGGGGGGGGATTTCTCCCTGGGCTGCTGCTGGGGCAAGGGACC 60 | | | | | | Qy | 951 | |
| Db | 16 ATGGGGCTGGGGGGGGGATTTCTCCCTGGGCTGCTGCTGGGGCAAGGGACC 75 | | | | | | Db | 1021 | CTGTCAGCTACAGAACACCAACTCTGGGAGGAGCTGAGCTGAGAAGAGCT |
| Qy | 61 CCTGAGATCATACTCTCTCTAGAAAGAGTACATGGAAAGGGCACTATGCCAGAGTAC 120 | | | | | | Qy | 1081 | |
| Db | 76 CCTGAGATCATACTCTCTCTAGAAAGAGTACATGGAAAGGGCACTATGCCAGAGTAC 133 | | | | | | Db | 952 | CTGTCAGCTACAGAACACCAACTCTGGGAGGAGCTGAGCTGAGAAGAGCT |
| Qy | 121 CTGAGAAATTCTCCATTCCATGCAACCAAATAATTAGGAGTAAATGCCCTCTAAA 180 | | | | | | Qy | 1011 | |
| Db | 134 ----- 133 | | | | | | Db | 1080 | CTGTCAGCTACAGAACACCAACTCTGGGAGGAGCTGAGCTGAGAAGAGCT |
| Qy | 181 TCACITCACTCTGAGACCTTCGATATAATTGGCAGGCTTGGCAATCTGGAA 240 | | | | | | Qy | 1012 | 1012 GAGGTAAAGAGACGGGAAATCATGAGTCTCTGAGCAAGAACCCCTCGATACGGTTC |
| Db | 134 -CACITCACTCTGAGACCTTCGATATAATTGGCAGGCTTGGCAATCTGGAA 192 | | | | | | Db | 1141 | ACCTCAGTTCTAGCTGAGGCCGCTGAGCATCAAGTAATTGGAGCCCTGGGACTCT |
| Qy | 241 ATTCCTCCAGACCCAGGGACGGGCGGATCTGGAGAGAGCAACTGGACTGAGCAAGGCCACAG 300 | | | | | | Qy | 1200 | |
| Db | 193 ATGGCCAGACCCAGGGACGGGCGGACATCTGGAGAGAGAC-----GAAAGCCACAG 246 | | | | | | Db | 1072 | ACCTCAGTTCTAGCTGAGCCGCGCTGAGCATCAAGTAATTGGAGCCCTGGGACTCT |
| Qy | 301 CAGCTGTCGATGGGAGTGTGAGCACTGCTCTGAGCATACATCAGATATTGAGCTCTCT 360 | | | | | | Qy | 1131 | |
| Db | 247 CAGGGTGCATGGGGGTGTGAGCATCTCTCTGAGCATCATCAGATATTGAGCTCTCT 306 | | | | | | Db | 1201 | CTCACGGCAGGCAATTGGCGGGGTCACCTGGGAGCTGGCTGGGACTCTGGAGCTGAGC |
| Qy | 361 GTTCGCAATGCGCTGGCGCAACTGGAGAGAGTCAATCCCAGGATGGCTGGAGAC 420 | | | | | | Qy | 1260 | |
| Db | 307 GTTCGCAATGCGCTGGCGCAACTGGAGAGAGTCAATCCCAGGATGGCTGGAGAC 366 | | | | | | Db | 1132 | AGTACCGAGGCCCTGCTCTGGAGGCGCTGGAGCTGGAGACCTCTGGAGCTGACTCAG |
| Qy | 421 TTGGTGGATTCAAGGCTCAAGACTGGAGAACATGGAGAGAGTCAATCCCAGGATGGCTGGAGAC 480 | | | | | | Qy | 1251 | |
| Db | 367 TTGGTGGATTCAAGGCTCAAGACTGGAGAACATGGAGAGAGTCAATCCCAGGATGGCTGGAGAC 423 | | | | | | Db | 1192 | TACCGAGGCCCTGCTCTGGAGGCGCTGGAGCTGGAGACCTCTGGAGCTGACTCAG |
| Qy | 481 CAATTGACTGGAAACTCATCAATGTTCTCACTAATGCAACGAGTGTACCTGTC 540 | | | | | | Qy | 1321 | |
| Db | 424 CAATTGACTGGAAACTCATCAATGTTCTCACTAATGCAACGAGTGTACCTGTC 483 | | | | | | Db | 1312 | GGGACACATCTCCCGGAATTCAACCTTCTGAGGGCTCTGGAGCTGGAGCTGGAGCTGGAA |
| Qy | 541 CCCTCTGCTAACAGAAATGGGCTTGGGGGGGGGGTGTGATGAGCTGGCTGAGTGTCAAC 1600 | | | | | | Qy | 1311 | |
| | | | | | | | Db | 1381 | GAACACAGTCCTAAACCGGCTCTTAAACCGGCTGAGGAGCTGGAGCTGGAGCTGGAGCTGGAA |
| | | | | | | | Qy | 1440 | |
| | | | | | | | Db | 1312 | GAACACAGTCCTAAACCGGCTGAGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAA |
| | | | | | | | Qy | 1365 | |
| | | | | | | | Db | 1441 | CTGTGCCAGGCCAGGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAA |
| | | | | | | | Qy | 1500 | |
| | | | | | | | Db | 1366 | -----CAGCCAGCTGCTCTTAAACCGGCTGAGGAGCTGGAGCTGGAGCTGGAGCTGGAA |
| | | | | | | | Qy | 1419 | |
| | | | | | | | Db | 1561 | GATCTGGCTCACTATCTCCAGGACTTCAGCTGAGCTGGAGCTGGAGCTGGAGCTGGAA |
| | | | | | | | Qy | 1540 | |
| | | | | | | | Db | 1480 | GATCTGGTAACTATCTCCAGGACTTCAGCTGAGCTGGAGCTGGAGCTGGAGCTGGAA |
| | | | | | | | Qy | 1539 | |
| | | | | | | | Db | 1420 | GAAGACTGGAGATAATAACCTGTTATAGGCGCAATGACCTGAGCTGGAGCTGGAA |
| | | | | | | | Qy | 1479 | |
| | | | | | | | Db | 1621 | CTCCATGCTGAGGTTCTGGCATTTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAA |

| | | | | | |
|----|--|------|----|--|------|
| Db | 1540 CTCATGCTGGTTCTCGGCATTGTCACTGGTCAAGTGTCAC | 1599 | Qy | 2758 CCAGTCGCGCAGGCCAGGTTTGTGTAACCTCGTCAGCTGTCAC | 2817 |
| Qy | 1681 CTGGGGACTGTACCTGAGGAAAGCTACTGCCCAGGATGATCCTCGTC | 1740 | Db | 2680 CCAGTCGCGCAGGCCAGGTTTGTGTAACCTCGTCAGCTGTCAC | 2739 |
| Db | 1600 CTGAGGAGCTGTACCTGAGGAAAGCTACTGCCCAGGATGATCCTCGTC | 1659 | Qy | 2818 GACSTASCGAGGTGGAGGAGCTTACGGGAGGCTACCGGAGAAGCATGCGCAGTC | 2877 |
| Qy | 1741 TGTCCCTGCTCTGAGTTGATACTAACAGACTTGCTACCTCTGGAATC | 1800 | Db | 2740 GAGCTAGCAGGTTGGAGGCTTACGGGAGGCTACCGGAGAAGCATGCGCAGTC | 2799 |
| Db | 1660 TGTCCCTGCTCTGAGTTGATACTAACAGACTTGCTACCTCTGGAATC | 1719 | Qy | 2878 GGGTCAAGCCCTATGAGACGGGAGGAGCTTCCTGGCTCGAGGAGGCTTCAG | 2937 |
| Qy | 1801 AACAGAAACTTCAGGAGAGACCACACAGTGTAGAGATGGGGAATGAC | 1860 | Db | 2810 GSCTCAGGCCGCTATGAGACGGGAGGAGCTTCCTGGCTCGAGGAGGCTTCAG | 2859 |
| Db | 1720 AACAGAAACTTCAGGAGAGACCACACAGTGTAGAGATGGGGAATGAC | 1779 | Qy | 2938 AACATCCAGCTCTGCTCTGGGGATGGGCTCCAGATACTGCTCTTGGCCAGAC | 2997 |
| Qy | 1861 GAAGATTACTGTTGCTGAGCGTCTTGAAAGCTGACATCCAGGACTCG | 1920 | Db | 2860 AACATCCAGCTCTGCTCTGGGGATGGGCTCCAGATACTGCTCTTGGCCAGAC | 2919 |
| Db | 1780 GAAGATTACTGTTGCTGAGCGTCTTGAAAGCTGACATCCAGGACTCG | 1839 | Qy | 2938 TGCATCCCAAACTAGAGAAATTCAGAAATTCCACTCCAGCTGGCAGAGCT | 3057 |
| Qy | 1921 GAAGATTACTGTTGCTGAGCGTCTTGAAAGCTGACATCCAGGACTCG | 1980 | Db | 2920 TGCATCCCAAACTAGAGAAATTCAGAAATTCCACTCCAGCTGGCAGAGCT | 2979 |
| Db | 1840 GAAGATTACTGTTGCTGAGCGTCTTGAAAGCTGACATCCAGGACTCG | 1899 | Qy | 3058 CTGACCACTTGAGGAAACAGAGGAGGAGGCTTGAGGAGGAGGAGCT | 3117 |
| Qy | 1981 CACTGCCGAGCAGCCAGTCTCTTCTCGCTCTGAGACATATGCTGAGAGG | 2040 | Db | 2990 CTGACCACTTGAGGAAACAGAGGAGGAGGCTTGAGGAGGAGGAGCT | 3039 |
| Db | 1900 CACTGCCGAGCAGCCAGTCTCTTCTCGCTCTGAGACATATGCTGAGAGG | 1959 | Qy | 3118 TGTCCTACTCAGATGAGGCCCTCTGAGAACCCCTGGAAATGAACTACAGTACCC | 3177 |
| Qy | 2041 ACTGCTCATAGTTGAAACAGATCAATATCACATGTCGAGGAGTCAC | 2097 | Db | 3040 TGTCCTACTCAGATGAGGCCCTCTGAGAACCCCTGGAAATGAACTACAGTACCC | 3099 |
| Db | 1960 ACTGCTCATAGTTGAAACAGATCAATATCACATGTCGAGGAGTCAC | 2019 | Qy | 3118 ATCAGGCAAGCTATGAGAACTTGAGGACTCGGAGTGTACTCTGTCAGAGTGGAGGT | 3237 |
| Qy | 2098 TTCTGAGACTACAAGAACATGAGGGCATGGACCTCTCCAGGAGG | 2157 | Db | 3100 ATCAAGCAGCATGGAGAACTGGGAGTGTACTCTGTCAGAGTGGAGGT | 3159 |
| Db | 2020 TTCTGAGACTACAAGAACATGAGGGCATGGAGGACTCTGTCAGAGTGGAGG | 2079 | Qy | 3218 ATAGTGTCCAACTCTGTCACCAACCTCCGACACCAGAGCATCAAGTGTGGGCC | 3297 |
| Qy | 2158 AGAGCCCTCTGCTGCTGACCTACTCTGAGCTGAGGCTGAGACATCAA | 2217 | Db | 3160 AATAGTGTCCAACTCTGTCACCAACAGAGGACCTCCGACCC | 3219 |
| Db | 2080 AGAGCCCTCTGCTGACCTACTCTGAGCTGAGGCTGAGACATCAA | 2139 | Qy | 3298 CTGGGTGACTCTGACTACAGGGAGCTGGAGGAGTCGACCAAACACTCCAGTACCC | 3357 |
| Qy | 2218 GTTGTGGCTCTGGGAGATCTGACCGCTGCAATGGAAATGGCTCCAAACAGAC | 2277 | Db | 3220 CTGGGTGACTCTGACTCTGACTCTGTCAGGCTGGAGGAGTCGACCAAACACTCCAGTACCC | 3279 |
| Db | 2140 GTTGTGGCTCTGGGAGATCTGACCGCTGCAATGGCTCCAAACAGAC | 2199 | Qy | 3358 ACATCTTGAGGGACTCTCTGGAGCATGGGGAGTGGAACTTGAGGACTCACCC | 3417 |
| Qy | 2278 GACCTCCCGATGTCACCAACAGTATGGGACTCTGAGGGAGGAGGCC | 2337 | Db | 3280 ACATCTTGAGGGACTCTCTGGAGCATGGGGAGTGGAACTTGAGGACTCACCC | 3339 |
| Db | 2200 GACCTCCCGATGTCACCAACAGTATGGGACTCTGAGGGAGGCC | 2259 | Qy | 3418 ACATCTGCAACATTCTGAGAGGAGTCAACCTTACCTCTGTCGTTCTACACCA | 3477 |
| Qy | 2338 TCCCTGGAGATGTGACCAACCTTACCTATCTCGGAGTTACAGAACTCA | 2397 | Db | 3310 ACATCTGCGCACATTCTGAGAGGAGTCAACCTTACCTCTGTCGTTCTACACCA | 3399 |
| Db | 2260 TCCCTGGAGATGTGACCAACCTTACCTATCTCGGAGTTACAGAACTCA | 2319 | Qy | 3478 TGGAGGGAGCAGGAGGAGTAAATGTCGAGGGAGGGCAGAGCTGGACATCCA | 3537 |
| Qy | 2398 GGCTACCGCGTGGCACCGGTGATGCCATGAGCAGTGTCAATCAGCTT | 2457 | Db | 3410 TGCGAGGAGCAGCAGCACTATAGTCGAGGAGTCAGGAGTCAGGAGTCAGCCA | 3459 |
| Db | 2320 GCTTAGCGCGTGGCACGGGTGATGCCATGAGCAGGAGTCAATCAGCTT | 2379 | Qy | 3558 GCCCAGGGCTGGGAGCTCTGGAGTAGGCGATGAAAACAGGCCCTGCTTCAACCTGGAG | 3594 |
| Qy | 2458 CGGGAGCAAGGCTGAGATCTATGAGCCAGTGTCACTGAGGATGAA | 2517 | Db | 3460 GCCCAGGGCTGGGAGCTCTGGAGTAGGCGATGAAAACAGGCCCTGCTTCAACCTGGAG | 3519 |
| Db | 2380 CCCGAGCAAGGCTAGGATCTATGAGCCAGTGTCACTGAGGAGTCAA | 2439 | Qy | 3555 AAAGACTGAGCTGTCACACTCTTCTGGGTCAGGACTTGTCATTAATGAGG | 3654 |
| Qy | 2518 GATGATCATAGATAATTCTGAGAGCTGGAGGTCATCAAGTGTGGAGGC | 2577 | Db | 3550 AAAGACTGAGCTGTCACACTCTTCTGGGTCAGGACTTGTCATTAATGAGG | 3579 |
| Db | 2440 GATGATCATAGATAATTCTGAGAGCTGGAGGTCATCAAGTGTGGAGGC | 2499 | Qy | 3655 AAATCGGGGGCCACTCTGGCACGGAAATGTCAGGACATCCACAGGCCCTGGACATC | 3714 |
| Qy | 2578 AGGGATTATGTTGACTCTGAGCAGATTCGAGCTGTATCTCGAGGACTTGTC | 2637 | Db | 3560 AAATCGGGAGC-----GAAATGTCAGGACATCCACAGGCCCTGGACATC | 3627 |
| Db | 2500 AGGGATTATGTTGACTCTGAGCAGATTCGAGCTGTATCTCGAGGACTTGTC | 2559 | Qy | 3715 CTCTCTGGGGAGCTCCAAAGGGCTTCTGTCAGGTCAGGAGCTGGCTGGCTAGC | 3774 |
| Qy | 2638 CACTCCCGAACTGCTGAGGAGTGTCAATGAGGAGTGTCACTGCTGCAACCTC | 2697 | Db | 3628 CTCTCTGGGGAGCTCCAAAGGGCTTCTGTCAGGAGGAGCTGGCTGGCTAGC | 3687 |
| Db | 2560 CACTCCCGAACTGCTGAGGAGTGTCAATGAGGAGTGTCACTGCTGCAACCTC | 2619 | Qy | 3775 CTCTCTGGGGAGCTCCAAAGGGCTTCTGTCAGGAGGAGCTGGCTGGCTAGC | 3834 |
| Qy | 2698 GTGGACTCTGAAACCCACTATCATGGGGAGGTCTCTGGAAACCCAGAOAGTC | 2757 | Db | 3688 CTCTACGGGCAAGGGGAATGTCAGGAGCTGCAAGAACACTGCACTG | 3747 |
| Db | 2620 GTGGACTCTGAAACCCACTATCATGGGGAGGTCTCTGGAAACCCAGAACAGTC | 2679 | | | |

| | | | | | |
|----|---|------|----|--|------|
| Db | 241 CCCCATCTTCACTACAGAACAGCAACTACCTGACCAGACTGCAGAACCCAAAGC | 300 | QY | 2113 AAGACAG-----CATGGGGTCATGGACCTGGCTGCCATGCAGGGAC | 2157 |
| QY | 1075 AASCCTG----- | | QY | 1092 AGTAAGAAA----- | 1092 |
| Db | 301 AACGTTGGAGAAATCATTCTTATTCTTGTGTCTCATTTCAGTAAAGAGAA | 360 | Db | 1381 AGATTCTGGATTGTCTTATGCCCTTGGGTCTATGGACCTGGCTCCATCGGGAC | 1440 |
| QY | 1093 GACCGGAATCAGATGCTTGTACAAAGACCCCTCCGATACGGTCCACCTCAGTCA | 1152 | QY | 2158 AGAGCCCCCTTCGCTTGACCTTACCTTCAGTCATGCCCTGAGACCTGAGACATCAA | 2217 |
| Db | 361 GAGCGGAATCACATGCTCTGACAAAGACCCCTCCGATACGGTCCACCTCAGTCA | 420 | Db | 1441 AGAGCCCCCTTCGCTTGACCTTACCTTCAGTCATGCCCTGAGACCTGAGACATCAA | 1500 |
| QY | 1153 AGGCTGAAGCCGCTGACATCAAGTAATGAGCCCTCGGGTACGGTCTCTCACCGGAGC | 1212 | QY | 2219 GTTGTGGTCCTGGGGATTCCTGAGCGTGGCATGGATTGCCCAAACCGAC | 2277 |
| Db | 421 AGCTGAAGCCGCTGACATCAAGTAATGAGCCCTCGGGTACGGTCTCTCACCGGAGC | 480 | Db | 1501 GTTGTGGTCCTGGGGATTCCTGAGCGTGGCATGGATTGCCCAAACCGAC | 1560 |
| QY | 1213 ATGGGGGGGTTGTCACACTGSGAACGCTTGTGCTTGTACTACCCAGGCTG | 1272 | QY | 2278 GACCTCCCGATGTCACCCACAGTCTGGGACTGTGCTACAGTGGAGGGACGGC | 2337 |
| Db | 481 ATGGGGGGGTTGTCACACTGSGAACGCTTGTGCTTGTACTACCCAGGCTG | 540 | Db | 151 GACCTCCCGATGTCACCCACAGTCTGGGACTGTGCTACAGTGGAGGGACGGC | 1620 |
| QY | 1333 CGGGAATTCAAACCTTCCCTGAAAGGGCTCTCTGTGGACCTGAAAGAACCGTCT | 1392 | QY | 2338 TCCCTGGAGAATGACCCACTTACCTATATCTCTGGGACTTTACAGAAACCTGCT | 2397 |
| Db | 601 CGGAAATTCAAACCTTCCCTGAAAGGGCTCTCTGTGGACTGGAAAGAACCGTCT | 660 | Db | 1621 TCCTGGAGAATGACCCACTTACCTATATCTCTGGGACTTTACAGAAACCTGCT | 1680 |
| QY | 1393 ATGCCCTTCTTAAACCAGSCTGTGCAAGGGCGGAGCTGAGGATCTACCTTCCAGGCC | 1452 | QY | 2339 GGCTACCGGTGGCACCGGTGATGCCATGACAGAAGATCCCTAATGGGAGTTAACAGAAACCTGCT | 2457 |
| Db | 661 ATGCCCTTCTTAAACCAGGCTGTGCGAGGAGCTGAGGATCTACCTTCCAGGCC | 600 | QY | 1631 GSCCTACSGCTGGCACGGGATGCAATGACGAACTGCTTCAAGTCAAGTCA | 1740 |
| QY | 1453 AGAGGCTGCTGACTCTGAGAACTGACAAGGAACTTACCTTCAGAAAGACTGGAG | 1512 | QY | 2458 CCCGAGAAAGCTGGAGATCTATGGCCAGTCCAACTCTGTCAGAGATGAA | 2517 |
| Db | 721 AGGAGGCTGTTGGACTGTGATGAGAACTGACAAGGATACATTCAAGAAACTGGAG | 780 | Db | 1711 CCGGAGAAAGCTGGAGATCTATGGCCAGTCCAACTCTGTCAGAGATGAA | 1800 |
| QY | 1513 ATAATAACCTGTTATAGGGGAATGACCTCTGATGTTCTGCAATGATCTGGTCCAC | 1572 | QY | 2518 GATGATCTAGACTTAATTCCATGAGAACTGAGACTGAACTCATCACAGTCGAGG | 2577 |
| Db | 781 ATAATAACCTGTTATAGGGGAATGACCTCTGATGTTCTGCAATGATCTGGTCCAC | 840 | Db | 1801 GATGATCTAGACTTAATTCCATGAGAACTGAGACTGAACTCATCACAGTCGAGG | 1860 |
| QY | 1573 TATTCTCCCAGAACATTCAAGACAACTGAAAGGGCTTGGACATCTCTCATGTCAG | 1632 | QY | 2578 AGGAGTATGACTACTGACAGATCTGAACTGAGACTGAACTCATCACAGTCGAGG | 2637 |
| Db | 841 TATCTCCCAGAACATTCAAGACAACTGAAAGGGCTTGGACATCTCTCATGTCAG | 900 | QY | 1881 AGGGATTATGACTACTGACAGATCTGAACTGAGACTGAACTCATCACAGTCGAGG | 1920 |
| QY | 1633 GTTCTCTGGCATGTGAACTCTGAACTGAACTGAACTGAACTGAACTGAACTGAA | 1692 | Db | 2638 CATCTCCCAATGCTTGACGCTCTGACATGAGGGCTCAGAGTCCTGGTCACCTC | 2697 |
| Db | 901 GTTCTCTGGCATGTGAACTCTGAACTGAACTGAACTGAACTGAACTGAACTGAA | 960 | QY | 1921 CATCTCCCAATGCTTGACGCTCTGACATGAGGGCTCAGAGTCCTGGTCACCTC | 1980 |
| QY | 1693 TACCAAGGAAAMAAGTCACTGCCAACGATGATCCTCTAGGTCCTGGCCCTGTC | 1752 | QY | 2638 GTGACTCTCTGAACTCATGAGGGCTGTTGTTGTTGTTGTTGTTGAGAACCCAGAACAGTC | 2757 |
| Db | 961 TACCAAGGAAAMAAGTCACTGCCAACGATGATCCTCTAGGTCCTGGCCCTGTC | 1020 | Db | 1981 GTGAGCTCTGAACTCATGAGGGCTGTTGTTGTTGTTGTTGTTGAGAACCCAGAACAGTC | 2040 |
| QY | 1753 CTGAGTTGATGATACTCAACAGAACCTGCTACCCCTCATGAACTGAAAGATT | 1812 | QY | 2718 CCAGTCGCAGCCACACTATCATGGGAGGGTGTCTCTGGAAACCCAGAACAGTC | 2817 |
| Db | 1021 CTGAGTTGATGATACTCAACAGAACCTGCTACCCCTCATGAACTGAAAGATT | 1080 | QY | 2031 CCAGTCGCAGCCACGCTGGTGTGTTGTTGTTGTTGTTGAGAACCCAGAACAGTC | 2100 |
| QY | 1813 CAGCAGAACCCACCAACTGATGAGACTGCGCGATGACAAAGGAAGATTACT | 1872 | QY | 2818 GAGCTAGCCAGGTGGAGGCCTOAGGGACCTACGGAGGAGCATGGCGAGCTGT | 2877 |
| Db | 1081 CAGCAGAACCCACCAACTGATGAGACTGCGCGATGACAAAGGAAGATTACT | 1140 | Db | 2101 GAGCTAGCCAGGTGGAGGCCTACGGAGCTACGGAGCATGGCGAGCTGTG | 2160 |
| QY | 1873 GTGGTTGTCGACGCCCTTGTGAAACGCTGGACATGCCAAAGACCTCGGAAGGGATGCT | 1932 | QY | 2818 GGGTCAGGCCGGCTPATGACACGCGAGGAGACTCTCTGGTCTGAGGCCCTCTTCAG | 2937 |
| Db | 1141 GTGGTTGTCGACGCCCTTGTGAAACGCTGGACATGCCAAAGACCTCGGAAGGGATGCT | 1200 | Db | 2161 GGGTCAGGCCGGCTPATGACACGCGAGGAGACTCTCTGGTCTGAGGCCCTCTTCAG | 2220 |
| QY | 1933 GACACTCTTCTGCTCTGACTGTTCCACTCTGACGAGCTCTCCGAGCA | 1992 | QY | 2938 AACATCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 2997 |
| Db | 1201 GACACTCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 1260 | Db | 2221 AACATCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 2280 |
| QY | 1993 GCGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 2052 | QY | 2998 TGATCCCAATCAGAACTTCACTCCCACTGACAGAGCCCTGGACCATATG | 3057 |
| Db | 1261 GCGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 1320 | Db | 2341 CTGAAACCTGAGATGAGAAACAGAGACCCCTGGACCTGAGAGCATGCCCATCC | 2400 |
| QY | 2053 TTGAAACGAGATCAATCACATGTCGACCGAGGTCGACGCCCTTCAGTCAACCTG | 2112 | QY | 3118 TGGCCACTGAGATGAGAAACAGAGACCCCTGGACCTGAGAGCATGCCCATCC | 3177 |
| Db | 1321 TTGAAACGAGATCAATCACATGTCGACCGAGGTCGACGCCCTTCAGTCAACCTG | 1380 | Db | 2441 TGGCCACTGAGACCCCTGGACCTGAGAGCATGCCCATCC | 2460 |

| | | | | | |
|----|---|------|----|---|------|
| QY | ATCAAGCCATTTGAGAAGCTGGGAGTACTCTGTACAGTGAAAGGTTC | 3237 | Db | 3541 GTCAGGAGGGCGGGAGATCTTCATAATGAGCCTGGCACTGCGGAGTGGCTCTGGAGG | 3600 |
| Db | 2461 ATCAAGCCATTTGAGAAGCTGGGAGTACTCTGTACAGTGAAAGGTTC | 2520 | QY | 4318 TGCAGGAGGGCGGGAGATCTTCATAATGAGCCTGGCACTGCGGACT | 4365 |
| QY | 3238 ATAGTGTCCAACTCTGTGCCACGCCGCTGAGCAGCATCAGTGTTGGCC | 3297 | Db | 3601 TCAGGAGGGCGGGAGGAGATCTTCATAATGAGCCTGGCACT | 3648 |
| Db | 2521 ATAGTGTCCAACTCTGTGCCACGCCGCTGAGCAGCATCAGTGTTGGCC | 2580 | QY | 3298 CTGGGGACTCTGTACTACAGCAGCTGAGCTGAGCAACAACTCCAGTACCTTACCC | 3357 |
| Db | 2581 CTGGGGACTCTGTACTACCTGAGTAAGGGAGCTGAGCAACAACTCCAGTACCTTACCC | 2640 | QY | 3358 ACATCTGGAGGGACTCTGTAGGACATGGAGGAGTGGAACTTGAGACTCAACC | 3417 |
| Db | 2641 ACATCTGGAGGGACTCTGTAGGACATGGAGGAGTGGAACTTGAGACTCAACC | 2700 | QY | 3418 ACATCTGGCCAACTCTGTAGAAGAGTTAACCTTACCTCTGTGCTTACAGCACCC | 3477 |
| Db | 2761 TGGAGGGACAGCAGCAGACTAAATGTTGAGGACTACGGAGATGAGACTGACC | 2820 | QY | 3478 TGGAGGGACAGCAGCAGACTAAATGTTGAGGACTACGGAGATGAGACTGACC | 3537 |
| Db | 2701 ACATCTGGCCAACTCTGTAGAAGAGTTAACCTTACCTCTGTGCTTACAGCACCC | 2760 | QY | 3538 GCCCXGGCTGGACCTGGTAGGAGGAAAGCCCCGACATCAACTGGAGAA | 3597 |
| Db | 2821 GCCXGGCTGGACCTGGTAGGAGGAAAGCCCCGACATCAACTGGAGAA | 2880 | QY | 3598 GACTGGAGCTGGTCACCTCTCATGGGTCACGACTTGTTGTTACTGTGAAT | 3657 |
| Db | 2881 GACTGGAGCTGGTCACCTCTCATGGGTCACGACTTGTTGTTACTGTGAAT | 2940 | QY | 3658 CCGGAGGCCACTTCCACCGAATATGTCAGCACATCCACAGGCCCTGACATCTC | 3717 |
| Db | 2941 CCGGAGGCCACTTCCACCGAATATGTCAGCACATCCACAGGCCCTGACATCTC | 3000 | QY | 3718 TCTGAGGAGCTCCAAAGGGCTTCTGTCACCTGGTGGAGGTATGGAGCTGCTAGGCTG | 3777 |
| Db | 3001 TCTGAGGAGCTCCAAAGGGCTTCTGTCACCTGGTGGAGGTATGGAGCTGCTAGGCTG | 3060 | QY | 3778 TACCAAGGCCAAAGGGGAATGTGCATCTGGCGCTGAGAACAGCTGACTGCTC | 3837 |
| Db | 3061 TACCAAGGCCAAAGGGGAATGTGCATCTGGCGCTGAGAACAGCTGACTGCTC | 3120 | QY | 3838 AGACACTCGCAAAGCTCCCTGAGAAGCAACTGAGAAGACTGAACTCCAG | 3897 |
| Db | 3121 AGACACTCGCAAAGCTCCCTGAGAAGCAACTGAGAAGACTGAACTCCAG | 3180 | QY | 3898 CATGGGATCTCCAGTTCTCTACTCTGACCAACATAACACAGCGTGTAGGACTTGGGTT | 3957 |
| Db | 3181 CATGGGATCTCCAGTTCTCTACTCTGACCAACATAACACAGCGTGTAGGACTTGGGTT | 3240 | QY | 3958 GTGGTCAAGCTTCTCCAAACACACTACCCCTGTGACGAGGAGGGGACACTGAC | 4017 |
| Db | 3241 GTGGTCAAGCTTCTCCAAACACACTACCCCTGTGACGAGGAGGGACACTGAC | 3300 | QY | 4018 CTCACTCTCTGGAGACTGTCTCACTCTGAGCCGGCATGGAGATGGCC | 4077 |
| Db | 3301 CTCACTCTCTGGAGACTGTCTCACTCTGAGCCGGCATGGAGATGGCC | 3360 | QY | 4078 ATGGCACTCTGAGAACACATCTGAGGACCAACTGGCGCGAGACTACCTCAACACTC | 4137 |
| Db | 3361 ATGGCACTCTGAGAACACATCTGAGGACCAACTGGCGCGAGACTACCTCAACACTC | 3420 | QY | 4138 ACCCACAGGCCAGCAACTCAAGTGCCCTCTCTGGAGGACCTTACCTCTACCCCTG | 4197 |
| Db | 3421 ACCCACAGGCCAGCAACTCAAGTGCCCTCTCTGGAGGACCTTACCTCTACCCCTG | 3480 | QY | 4198 CGAACACGGCAATGCTCCAGGAGGGCTGAGAAGAGCCCCGAGGGCTACTCTGGCT | 4257 |
| Db | 3481 CGAACACGGCAATGCTCCAGGAGGGCTGAGAAGAGCCCCGAGGGCTACTCTGGCT | 3540 | QY | 4258 GTCCCGAGTGGAGGGAGTGGCCCTGTGTGGGGCATCTGGGAGAAGCTGTTGAGG | 4317 |

PR 10-JUL-2001; 2001US-304354P.
 PR 31-JUL-2001; 2001US-309198P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-3125430P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-33301P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-332184P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-009290.
 XX (CURA-) CURAGEN CORP.
 XX PT Padigaru M, Spyrek KA, Shenoy SG, Taupier RJ, Pena CBA, Li L;
 PI Zethusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Parturajan M, Ganguli E, Vernet CAM, Guo X, Tchernev V;
 PI Fernandes ER, Caenan SJ, Malvanykar UM, Gerlach V, Liu Y;
 PI Anderson D, Spaderna SK, Catterton B, Burgess C, Leite M, Zhong H;
 PI Alisbrook JP, Lepley DM, Rieger DK;
 XX DR; 2002-723332/78.
 XX DR-P-PSDB; ABUK5081.
 XX PT NOVX polypeptides and poly nucleotides, useful for preventing or
 PT treating a disorder associated with aberrant NOVX expression or
 PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
 PT bronchial asthma.
 XX PS Claim 13; Page 167-168; 1103pp; English.
 CC This invention describes novel human NOVX polypeptides which have
 CC hypotensive activity. Pharmaceutical compositions comprising the NOVX
 CC proteins or nucleic acid molecules or NOVX antibodies are useful for
 CC preventing or treating a disorder associated with aberrant NOVX
 CC expression or activity e.g. cancer, hypertension, atherosclerosis,
 CC cardiomyopathy or bronchial asthma. The products of the invention can
 CC be used for gene therapy or a vaccine. ABX97008-ABX97185 are cDNA
 CC fragments amplified and isolated by the PCR primers and probes
 CC represented in ABX13460-ABX13462 and ABX97186-ABX97593. ABX97008-ABX97185
 CC encode the NOVX proteins described in ABU65041-ABU65218.
 XX SQ Sequence 4268 BP; 1039 A; 1209 C; 1119 G; 901 T; 0 other;
 Query Match 75.2%; Score 3335; DB 24; Length 4268;
 Best Local Similarity 97.0%; Pred. No. 0; Mismatches 50; Indels 57; Gaps 9;
 Matches 3512; Conservative 0;
 OY 771 TAGGAAAGCTGGACAGCCTCCCTGGCCCTGGAGGAGTAGTGAGCAGGAGCTTCAC 830
 Db 687 TAGGAAAGCTGGACAGCCTCCCTGGCCCTGGAGGAGTAGTGAGCAGGAGCTTCAC 746
 OY 831 CGTGTGTTTCAAGCTTCTTCTATGAGACCCCATCTCACACTCGGAGAACCCCG 890
 Db 747 CGTGTGTTTCAAGCTTCTATGAGACCCCATCTCACACTCGGAGAACCCCG 794
 OY 891 ATCCAGGATTCTACACCTGGCTGGCATCTGGATAAGGTGATGAGCTGGAGCAGCAGGG 950
 Db 795 ATCCAGGATTCTACACCTGGCTGGCATCTGGATAAGGTGATGAGCTGGAGCAGGG 854
 OY 951 AGAGAAAGATGGCATGGTAAACACCGGAGGCCATAGAAGTGTCCCTCAGGA 1010
 Db 855 AGAGAAAGATGGCATGGTAAACACCGGAGGCCATAGAAGTGTCCCTCAGGA 914

OY 1011 GACCCCTATCTGTCAGCTACAGAAACAGCAACTACCTGACCACTGAGAACCCA 1070
 Db 915 GAGCCCTPATCTGTCAGCTACAGAAACAGCAACTACCTGACCACTGAGAACCCA 974
 OY 1071 AGCAAGCTTGAGTAAGAGAGAGAGGGAAATCAGTGTCTGCAGCAAGACCCCTCGA 1130
 Db 975 AGACAAGCTT--GTAAGAGAAGAGGGAAATCAGTGTCTGCAGCAAGACCCCTCGA 1031
 OY 1131 TAGGTTCCACCTCAGTCTAGAGCTAAGCGGCTACATCACCTTAATGGCC 1190
 Db 1032 TAGGTTCCACCTCAGTCTAGAGCTAAGCGGCTACATCACCTTAATGGCC 1091
 OY 1191 GGTTGACTCTCTACGGAGGATATGGGCGGGTCAACCTGGAGLGTCTGGAGGT 1250
 Db 1032 GGTTGACTCTCTACGGAGGAGATGGGCGGGTCAACCTGGAGLGTCTGGAGGT 1151
 OY 1251 CTGGACTCTAGTACCGAGGCTGTCCTAGGCTGAGCTGGAGATGAGACATGGACCGT 1310
 Db 1152 CTGACTCTAGTACCGAGGCTGTCCTAGGCTGAGCTGGAGATGAGACATGGACCGT 1211
 OY 1311 TACCACTGGGACATCTCGGAAATCAACCTTCTCTGAAGGCTCTCTGG 1370
 Db 1212 TACCACTGGGACATCTCGGAAATCAACCTTCTCTGAAGGCTCTCTGG 1271
 OY 1371 CACTGGAAAGAACCTCTTAACCCAGGCTGAGGGCCGGC 1430
 Db 1272 CACTGGAAAGAACCTCTTAACCCAGGCTGAGGGCCGGC 1331
 OY 1431 TGAGGATCTACCTGCTGGCCAGGAGCTGGTGGACCTGTGAGAATGACAGGAT 1490
 Db 1332 TGAG-----CHGCCAGGAGCTGGTGGACCTGTGAGAATGACAGGAT 1379
 OY 1491 ACACTTTACAGGAGCTGGAGTAATACCCGTTATAGCGGCAATGACTCTG 1550
 Db 1380 ACACTTTACAGGAGCTGGAGTAATACCCGTTATAGCGGCAATGACTCTG 1439
 OY 1551 TTCTGCAATGATCTGGCCACTATCTCCCAGAACTCACAGACACATTGGAAGGC 1610
 Db 1440 TTCTGCAATGATCTGGCCACTATCTCCCAGAACTCACAGACACATTGGAAGGC 1499
 OY 1611 CCTCGACATCTCTCATCTG-----AGTGTCTCGGCAATTGGAAGCTGAGGT 1664
 Db 1500 CCTCGACATCTCTCATCTG-----AGTGTCTCGGCAATTGGAAGCTGAGGT 1559
 OY 1655 GCTTGAGATGTCGAACTTGAGGGAGCTTACCGAGGAGAAAGTCTATGCCAAGAT 1724
 Db 1500 GCTTGAGATGTCGAACTTGAGGGAGCTTACCGAGGAGAAAGTCTATGCCAAGAT 1619
 OY 1725 GATCTCTCGGTCTGTCCTGGCTGCTGAGTTGATGATACTCACAGACTGC 1784
 Db 1620 GACCTCTCGGTCTGTCCTGGCTGCTGAGTTGATGATACTCACAGACTGC 1679
 OY 1735 TACCTCTCTGAAATTCAAAGAGTTTCAAGGAGAACCCCAACTATGAGAGCTGG 1844
 Db 1680 TACCTCTCTGAAATTCAAAGAGTTTCAAGGAGAACCCCAACTATGAGAGCTGG 1739
 OY 1815 GCGATATGACACAGGGAGATTACTGTGTTGAGGAGAACCCCAACTATGAGAGCTGG 1904
 Db 1740 GCGATATGACACAGGGAGATTACTGTGTTGAGGAGAACCCCAACTATGAGAGCTGG 1799
 OY 1905 CATGCCAAGACCTCGGAGGATGCTGACACTCTTCTGCTCTGACTCTGTTCCA 1964
 Db 1800 CATECCAAGACCGAGGAGATGCTGACACTCTTCTGCTCTGACTCTGTTCCA 1859
 OY 1955 CTTCAGCGCAAGCTCTCTCCGGAGRCGCACTGCTCTGAAACATATCTGGAGCC 2024
 Db 1860 CTTCAGCGCAAGCTCTCCGGAGRCGCACTGCTCTGACTCTGTTCCA 1919
 OY 2025 TGTGGCAGAGAGCTGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2084
 Db 1920 TGTGGCAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1979

| | | | | | | | |
|----|------|---|------|----|------|--|------|
| QY | 2085 | CAGAGTCCAGCGGTTCTGAGGACCTACAGAACAGCATGCAGCAGCTATGGACCTGTCT | 2144 | Db | 3050 | TAACTACAGTACCCATCAAGCAGCATGAGAAGCTGGGAGTGAGCTCTGTGAC | 3119 |
| Db | 1980 | ACAGTCAGCGGTTCTGAGGACCTACAGAACAGCATGCAGCAGCTATGGACCTGTCT | 2039 | QY | 3222 | AGAGTGGAGGCTCCAAATAGTTCCAACTCTGTGCCACCTCTGCCACCAAGAGAT | 3281 |
| QY | 2145 | GCCATGCAGGGACAGAGCCCTTCGCTCTGCTGCAACCTACCTCATGATGCCCTGAGACC | 2204 | Db | 3120 | AGAGTGGAGGCTCCAAATAGTTCCAACTCTGTGCCACCAAGAGAT | 3179 |
| Db | 2040 | GCCATGCAGGGACAGAGCCCTTCGCTCTGCTGCAACCTACCTCATGATGCCCTGAGACC | 2099 | QY | 3282 | CAAAGTGGGGGCCCTGGGACTCTCTGACTACAGCAGTGGAGCTCGACCAAAAGAA | 3341 |
| QY | 2205 | TGAGACATCCAACTGTGGCTGCTGCTGGGGATCTCTGACCCCTGAAATGAATTGG | 2264 | Db | 3180 | CAAAGTGGGGGCCCTGGGACTCTCTGACTACAGCAGTGGAGCTCGACCAAAAGAA | 3239 |
| Db | 2100 | TGAGACATCCAACTGTGGCTGCTGCTGGGGATCTCTGACCCCTGAAATGAATTGG | 2159 | QY | 3312 | CTCCAGTACCTACCCACATCTTGAGGAGCTCTGGGACTCTGCTCACCTGGAGAA | 3401 |
| QY | 2265 | CTCCAAACAGACCTCCGCTGCTGCAACCTACAGATTCGGGACTCTGCTCACAGCT | 2324 | Db | 3240 | CTCCAGTACCTACCCACATCTTGAGGAGCTCTGGGACTCTGCTCACCTGGAGAA | 3299 |
| Db | 2160 | CTCCAAACAGACCTCCGCTGCTGCAACCTACAGATTCGGGACTCTGCTCACAGCT | 2219 | QY | 3402 | CTTGAGGACTCACCCACATCTTGAGGAGCTCTGGGACTCTGCTACAGCTACAGCT | 3461 |
| QY | 2325 | AGAGGGGAAGGCTCCCTGAGAAGATGTGACCCACCTTACCTA---ATATCTGGAGTT | 2381 | Db | 3310 | CTTGAGGACTCACCCACATCTTGAGGAGCTCTGGGACTCTGCTCACCTTACCTGGAGTT | 3359 |
| Db | 2220 | AGAGGGGAAGGCTCCCTGAGAAGATGTGACCCACCTTACCTA---ATATCTGGAGTT | 2279 | QY | 3462 | CTTCTCTTACAGCACCTGGGAGCTGGGACTCTGCTGAGGAGTTCACCCCTACCTCTGGAGTT | 3521 |
| QY | 2382 | TAACAGAAACCTACAGGCTAACGGCTAACGGCTGATGCCATGACGAGATGCTT | 2441 | Db | 3350 | CTTCTCTTACAGCACCTGGGAGCTGGGACTCTGCTGAGGAGTTCACCCCTACCTCTGGAGTT | 3638 |
| Db | 2280 | TAACAGAAACCTACAGGCTAACGGCTAACGGCTGATGCCATGACGAGATGCTT | 2339 | QY | 3519 | CGACATCAGCTGAGAAGACTGGGAGCTGGGACTCTGCTGAGGAGTTCACCCCTACCTCTGGAGTT | 3419 |
| QY | 2442 | CTCTAACTCAAGCTTCCGGACAAAGCTGGGACTCTGCTGAGGCTTCACCTCAACTCT | 2501 | Db | 3410 | -- CATACTTCAGGAGCTGAGGAGTAAATACCTGAGACTGAAAGTCAACCTTACCTCTGGAGTT | 3536 |
| Db | 2340 | CTCTAACTCAAGCTTCCGGACAAAGCTGGGACTCTGCTGAGGCTTCACCTCAACTCT | 2399 | QY | 3522 | AGCT---AGGAGCATGAGCCAGGCTGGGAGCTGGGACTCTGCTGAGGAGTTCACCCCTACCTCTGGAGTT | 3578 |
| QY | 2502 | GATGAGAAGATGAAAGATGATGATGATGAGATAAGTAATTCCATGAGACTGAGAAGTCAACATGAGT | 2561 | Db | 3420 | AGCTAGGAGGACATGCCAGCCAGGCTGGGACTCTGCTGAGGAGTTCACCCCTACCTCTGGAGTT | 3479 |
| Db | 2400 | GATGAGAAGATGAAAGATGATGATGAGATAAGTAATTCCATGAGACTGAGAAGTCAACATGAGT | 2459 | QY | 3559 | CGACATCAGCTGAGAAGACTGGGAGCTGGGACTCTGCTGAGGAGTTCACCCCTACCTCTGGAGTT | 3638 |
| QY | 2562 | AGTCTGATGGAGGAGGATPATGTCACACTTGACAGATGAGATCTSPATTCTGC | 2621 | Db | 3440 | -- CATACTTCAGGAGCTGAGGAGTAAATACCTGAGACTGAAAGTCAACCTTACCTCTGGAGTT | 3536 |
| Db | 2460 | AGTCTGATGGAGGAGGATPATGTCACACTTGACAGATGAGATCTSPATTCTGC | 2519 | QY | 3639 | ACGGCCCTGGAGCATCCCTCTGAGGAGCTCCAAGGGCTTCTGCAAGCTGGAGGT | 3758 |
| QY | 2622 | AGCCAACTTGTGACCCAACTCCCTGAGACCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGG | 2681 | Db | 3585 | ACGGCCCTGGAGCATCCCTCTGAGGAGCTCCAAAGGGCTTCTGCAAGCTGGAGGT | 3644 |
| Db | 2520 | AGCCAACTTGTGACCCAACTCCCTGAGACCTGGGAGCTGGGAGCTGGGAGCTGG | 2579 | QY | 3719 | CATGGAGGCTGCTAGCTGTACCGAGCTGCTGAGGAGCTGGGAGCTGGGAGCTGG | 3818 |
| QY | 2682 | AGTCTGGTCACCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCT | 2741 | Db | 3557 | CTGTGATTCTGCAATGTC-----TGGTAGGTGATATGTCAGCACATCA | 3584 |
| Db | 2580 | AGTCTGGTCACCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCT | 2639 | QY | 3659 | ACGGCCCTGGAGCATCCCTCTGAGGAGCTCCAAGGGCTTCTGCAAGCTGGAGGT | 3704 |
| QY | 2742 | AAACCCAGAACGAGCTGGCCAGGCTTGTGAGCTCTGAGCTCTGAGCTCTGAGCT | 2801 | Db | 3585 | ACGGCCCTGGAGCATCCCTCTGAGGAGCTCCAAAGGGCTTCTGCAAGCTGGAGGT | 3644 |
| Db | 2640 | AAACCCAGAACGAGCTGGCCAGGCTTGTGAGCTCTGAGCTCTGAGCTCTGAGCT | 2699 | QY | 3719 | CATGGAGGCTGCTAGCTGCTGAGGAGCTGCTGAGGAGCTGGGAGCTGGGAGCTGG | 3764 |
| QY | 2802 | GCAGGAGACTCCAAAGCTAGCAGGCTGAGGCTCTGAGGAGCTCTGAGCTCTGAGCT | 2861 | Db | 3615 | CATGGAGGCTGCTAGCTGCTGAGGAGCTGCTGAGGAGCTGGGAGCTGGGAGCTGG | 3704 |
| Db | 2700 | GCAGGAGACTCCAAAGCTAGCAGGCTGAGGAGCTCTGAGCTCTGAGCTCTGAGCT | 2759 | QY | 3819 | GAACAACGACTGAGCTGCTGAGGAGCTGCTGAGGAGCTGGGAGCTGGGAGCTGG | 3878 |
| QY | 2862 | CATGGAGGCTGAGGAGCTGAGGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCT | 2921 | Db | 3765 | AGTGAACGACTGAGCTGCTGAGGAGCTGCTGAGGAGCTGGGAGCTGGGAGCTGG | 3824 |
| Db | 2760 | CATGGAGGCTGAGGAGCTGAGGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCT | 2819 | QY | 3919 | GCCTGAGGACTTGGGGTGTGGAGCTGAGGAGCTGGGAGCTGGGAGCTGGGAGCTGG | 3998 |
| QY | 2922 | GCAGCCCTTCTGAGAACATCAGCTCTGAGGAGCTGAGCTCTGAGCTCTGAGCT | 2981 | Db | 3825 | GCCTGAGGACTTGGGGTGTGGAGCTGAGGAGCTGGGAGCTGGGAGCTGGGAGCTGG | 3884 |
| Db | 2820 | GCAGCCCTTCTGAGAACATCAGCTCTGAGGAGCTGAGCTCTGAGCTCTGAGCT | 2879 | QY | 4059 | CGGGCATCGGAGTGCACATCTGAGGAGCTGAGGAGCTGGGAGCTGGGAGCTGGGAGCTGG | 4118 |
| QY | 2982 | CTCTTGTGAGCTGAGCTGAGAACATCAGCTCTGAGGAGCTGAGCTCTGAGCT | 3041 | Db | 3942 | CGGGCATCGGAGTGCACATCTGAGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGG | 4001 |
| Db | 2880 | CTCTTGTGAGCTGAGAACATCAGCTCTGAGGAGCTGAGCTCTGAGCTCTGAGCT | 2939 | QY | 4119 | GACTACTCCAACTCTGAGGAGCTGAGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGG | 4178 |
| QY | 3042 | CTTGTGAGAACATCAGCTGAGGAGCTGAGGAGCTGAGCTCTGAGCTCTGAGCT | 3101 | Db | 4002 | GACTACCCAACTCTGAGGAGCTGAGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGG | 4061 |
| Db | 2940 | CTTGTGAGAACATCAGCTGAGGAGCTGAGGAGCTGAGCTCTGAGCTCTGAGCT | 2999 | QY | 4179 | CCCTTACCTCTACCCCTGGGAGACGGCTTCTGAGGAGCTGGGAGCTGGGAGCTGGGAGCTGG | 4238 |
| QY | 3102 | AGAGATGCCATCACCTGCTGAGGAGCTGAGGAGCTGAGCTCTGAGCTCTGAGCT | 3161 | Db | 4062 | TCTTACCTCTACCCCTGGGAGACGGCTTCTGAGGAGCTGGGAGCTGGGAGCTGGGAGCTGG | 4121 |
| Db | 3000 | AGAGATGCCATCACCTGCTGAGGAGCTGAGGAGCTGAGCTCTGAGCTCTGAGCT | 3059 | QY | 4239 | CGGGCTCTACTGAGGAGCTGAGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGG | 4298 |
| QY | 3162 | TAACTACAGTACCCATCAAGCCAGGCCATTGAGAACTGGGGAGTGTGACTCTGTGAGCT | 3221 | | | | |

Db 4122 CGAGGTGCTTACTGGCTTCCAGTCAGGAGTTGGCTTGTTGGCATCAT 4181
 Qy 4299 CGGCACTGGCTGGAGGGAGTGCAGGAGGTGCAGCGAGGAGATCCGCCATTGASCT 4358
 Db 4182 CGGACAGTGTGAGGTGGAGGGAGTGGCTCCTCCAAAGATCCCTCCAAAGAGCT 4241
 Qy 4359 GCGCACTGGCCCTCTAG 4377
 Db 4242 GCGCACTGGCCCTCTAG 4260

RESULT⁸

ID AAT91874 standard; cDNA; 4613 BP.
 XX
 AC AAT91874;
 XX
 DT 28-JAN-1998 (first entry)
 XX
 DE Rat phospholipase-B/lipase cDNA from plasmid pSVL-RPLB.
 XX
 KW phospholipase B; lipase; supplement; pancreatic phospholipase; reagent;
 KW screening; rat; db.
 XX
 OS Rattus rattus.
 FH location/Qualifiers
 FT CDS 35..4308 /*tag= a
 FT sig_peptide 36..125 /*tag= b
 FT mat_peptide 126..435 /*tag= c
 FT /note= "putative mature protein"
 PN JRP09248190-A.
 XX 22-SEP-1997.
 XX 15-MAR-1996; 96JP-0086022.
 PR 15-MAR-1996; 96JP-0086022.
 PA (TOJO/) TOJO H.
 PA (TORU) TONEN CORP.
 XX WPI; 1997-520744/48.
 DR P-PSDB; AAW0751.
 XX DNA encoding protein with phospholipase B and lipase activity - useful as supplement to pancreatic phospholipase
 PT Claim 4; Pages 7-12; 16pp; Japanese.

The present cDNA encodes a protein which has both phospholipase B and lipase activity. The claimed novel protein comprises at least amino acids 367-712 of AAW0751. The phospholipase is useful as a supplement to pancreatic phospholipase and as a reagent for the determination of phospholipase B/lipase.

SQ Sequence 4613 BP; 1218 A; 1243 C; 1142 G; 1010 T; 0 other;

Query Match 61.4%; Score 2687; DB 18; Length 4613;
 Best Local Similarity 77.1%; Pred No. 0; Mismatches 970; Indels 21; Gaps 5;
 Matches 3340; Conservative 0;

Qy 28 CTGGAGCTGCTGCTCTGGGCAAGGACCTCTAGATCCATACCTCTCTAGAAG 87
 Db 69 CTGCTGCTGCTGCTACTGCTGGGCAAGGACCTCTCCAAATCCAGGCTCTGGAGA 128
 Qy 88 AGTAGATGGAGGGAGGAGTATGGCCAGAGAACCTGAAGAAATCTCCATGCAAC 147
 Db 129 AACACATGGCCAGCCCCAGGAGTTCTGGAGCCCTGAGAATTTTCAATCCCTGCAAG 188

Qy 148 CCAATAATAATTAGGAGCTATGCTCTTAATCAGTCACTCTGAGGCTCTGAT 207
 Db 189 CCAGAAGAGTGTAGACTAGTGCTTCTAAGTCAGTCCTCTGAGACCTCTGAG 248
 Qy 208 ATTAAATTGTGTCAGCAATTGCAATTGGAAATTCTCCAGACCCAGGGAGGGGG 265
 Db 249 ATTAACCTGTGGAGCCATCGGCACCTAGAAACTCTCCAGCCCTGGCTCAGGGTG 308
 Qy 256 ---ATCTGAGAGAGAGACTTGACTGTAAG---GCCACAGCAGGTTGCAATGGAGTG 318
 Db 309 GTGACACATGGAGAACCTCAAGAACCTGAGAGAACCTGAGGAACTACAGATGTGCAATGAGATTC 368
 Qy 319 ATGACAGTCTTCAGACATCATCAGATTTCAGICCTCTGTCATGCTGTGTC 378
 Db 339 ATGACAGCCTTCAGATATCATCAGACATTCAACCTTCGTCATGCTCACCGT 428
 Qy 379 CACACTGGAAAGAGTGTATACCCAC---GATGGTGTGAAACTTGACTTGAACTG 435
 Db 429 TCTCTGGAGAGGTACTGCAGGCCACACTACTATGGCAGAAGACTTGAACTG 488
 Qy 436 CAGAACTGGTGAAGAACATGAAGAGAACCTGCAACTTGACTTCATTTACTG 495
 Db 439 AAAGAGCTGTGAGGACACTGAAAGAACACCCGAACTTGATGTTGAGGACTG 548
 Qy 496 CTCATCAATGTTCTCAGTAAATGCAAGCCAGTGTACTCTGCCCCTGCTCAAC 555
 Db 519 CTGTCATGCTGCTCAGATCATCAGACATTCAACCTTCGTCATGCTCACCG 608
 Qy 556 AATGGGCTTGCAGGGGGGGGGCTGGATGAGCTGATGGGGCTGGAACTACCTG 615
 Db 609 AAAGAGCTGTGAAGAACACATGGAGATGCTCAGGGGCTGTGATTAATGCAAG 668
 Qy 616 GAGSTCCCAAGCATTTGTAACCTGGACCTCTTGAGSTCTGTGAGAGSTCTCTGT 675
 Db 659 GAGSTCCCAAGAGCTTGTGAATTGGATTCAGGTTTGTGAGAGTTCAGG 728
 Qy 676 CAGTATCGGACTTGGCTAGCCCTGACCCCTGACAGAGGCCCTGTAATGTCAGAGGAGC 735
 Db 729 CAGCATCAGAGACTGCTTCAGTAACACAGGCTGCACTCTGCTTCAGGATA 788
 Qy 736 ACCGGCTGGCCAGGGCTGGATGAGCTGAGGGCTTACCGAGATTTGCAAGGATA 795
 Db 789 ACCGAGTATCCAGGCTGTCTGAGCTGAGGGCTTACCGAGATTTGCAAGGATA 848
 Qy 796 GCCTCCACAGGATCACGGAGGAGTCCITCACCAGGTTTCCAGGAGGCTGGACAGCCTCTG 855
 Db 849 GCTTCACAGGCTCAATGAGTGAAG 908
 Qy 856 GAGACCACTCTCACTGGAGACCCGACTTCCAGGATTCACCGCTGCC 915
 Db 909 GAG-----GTGAGACTACCTCTGGAGGGCTCGCCCAAGGATTCACCGCTGCC 962
 Qy 916 TGCTCATCTGGATAGGATGATGGAGCACTGGAGGAGAAGATGAGCCATGAGTGA 975
 Db 963 CTAGGAGTCTGGATAGGATGATGGAGCACTGGAGGAGAAGATGAGCCATGAGTGA 1022
 Qy 976 AACACGGAGGCAATGAGTGTCTCTGGAGGAGCCCTATGTTGAGCTAGA 1035
 Db 1023 GCGAGAGAATACATTAATGTCCTCTGGAGGAGCCCTATGTTGAGCTAGA 1082
 Qy 1036 AACAGCAACTACCTGACAGACTGCGAGAACCCAGAGCAAGCTGAGGTAAGAGGA 1095
 Db 1083 AATAGCACTACCTGAGGAGACAGCTGAACTGAGGAACTTCAGGAGA 1142
 Qy 1095 GCGAACTAGAGTCTGCTGACAAAGACCCCTCGATAGGGTTCCACCTCAGTCAATGG 1155
 Db 1143 ACAAAATTACCTGCTGAGACCCCTGAGACTCCATCCCACAGTCAGG 1202
 Qy 1156 CTGAGGGGTGACATCAACGAAATGGAGCCCTGAGCTCTCAGGAGAAT 1215
 Db 1203 CTGAGGGGTGACATCAACGAAATGGAGCCCTGAGCTCTCAGGAGA 1262

| | | | | | |
|----|---|------|----|---|------|
| Qy | 1216 GGGGCCGGTCAACCTGGAACTGCTTGAGCTGACTCAGTCGGCGCTTC | 1275 | Db | 2343 AACAGATCAGGAGCTCTACAGTCGTGTTGGGACAAGTCTTGAGATGTCAC | 2402 |
| Db | 1263 GGGCAGGGTCAGCCCTGGAAATGCTGAGTCATCACTGGACCTCTTC | 1322 | Qy | 2356 ACCTTACCTAATCCCTGGAGTTAACAGAACCTACAGCTGGCGACG | 2415 |
| Qy | 1276 TGGAGCGTGGGGAGATGAGACATCGCACGGTACACCTGGAGAACCT | 1335 | Db | 2403 ACCTTSCCAACATCCCTCGGAAATTATSGAAATCTACAGGCTACGTCGAAAC | 2462 |
| Db | 1323 TGGAGCTGGGGAGATGAGACATCGCACGGTACACCTGGAGAACCT | 1382 | Qy | 2416 GTGATGCCAATGACAGGAATCCTCATCAAGTGTCCCGGAGAAAGGTGAG | 2475 |
| Qy | 1336 GATTCAACCTTCCCTGAAGGACTCTCTGTGGACTGSGAAAGAACCT | 1395 | Db | 2463 GTGACCTCAACTCTGCAAGCGGTCTTACAGGTGTCCCTGGGCAAGGCTG | 2522 |
| Db | 1383 GATTCAACCTCTCTGAGAGCTGGACTCTGAGAAAGAACACTCCCGA | 1442 | Qy | 2476 GATCTTAGGCCAAGTCCAACCTCTGATGAGAAGATGAAAGATCATAGAATA | 2535 |
| Qy | 1396 GCTCTCTTAACACGGCTGCGAGGGCCAGCTGGAGTCTACCTGTCAGCCAG | 1455 | Db | 2523 AACCTTGAGCTGAGTCAGTCAGCTGAGCTGATGAGAACAGAGTGAC | 2582 |
| Db | 1443 GATCCCTTCAACAGGCGTACAGGACCAATCTGATGGCTTACGCGCAA | 1502 | Qy | 2536 TTCCATGAGACTGGAGGTCTACAGCTGATCGGGAGGATATTGTGACTAC | 2595 |
| Qy | 1456 AGCTGGGGAGCTGATGAGAAAGTGAAGACGAGGAAATACTTCAGGAAGATA | 1515 | Db | 2583 TTTCACCAAGACTGGAGGTCACTGAGTGGCTTACGCGCAATTCCTGCGCTT | 2642 |
| Db | 1503 AGCTGGGGAGCTGATGAGAAAGTGAAGACGAGGAAATAACTTCAGGAAGATA | 1562 | Qy | 2656 GACCTCTGATGAGAAGTGGAGTCACCCCTTGAGAGACTGGAGCTTC | 2715 |
| Qy | 1516 ATTAACCTGTTATAGCGGCAATGACCTCTGATGATGATCTGCTCCACTAT | 1575 | Db | 2703 GACATCCCTGATAGGGATACCCAGGCTGTCACCTGAGCTGAGTT | 2762 |
| Db | 1563 ATCACTGTGTTATAGAGGGCAACGACCTCTGCTCTGCGCTT | 1622 | Qy | 2716 ACTATGTCGGCAGGTTCTGGAAACCCAGAGCAGTGCCAGGAGGCC | 2775 |
| Qy | 1576 TCTCCCGAAGAACCTTACAGACATCGAACATCCCTCAGCTGAGTT | 1635 | Db | 2763 AGTATCATTCGCAACTCTGTCAGTGGCTTACGCGCAATTCCTGACCAT | 2822 |
| Db | 1623 TCTCCCGAACCTTACAGACATCGAACATCCCTCAGCTGAGTT | 1682 | Qy | 2776 GTTTGTGTAACTGCGTTCTGACCCCTGGGAAACCTCCAGAGCTGAGCAG | 2835 |
| Qy | 1636 CCTCGGGATTCTGTGAACTCTGCTGAGCTGTCACCTGAGGGAGCTAC | 1695 | Db | 2823 GTCAGGATTCGATCTACTCTGCAAGCCAAATTCTGACCATCTCCGCAATGCTG | 2882 |
| Db | 1683 CCCGGGCCTTGTGAGATGAGTACCCCTGCGCTTCTGCTGCT | 1742 | Qy | 2836 GCCTTCGGCAGCCACTGGAGCATGGGAGCTGTCACCTGTCACCTGACCCC | 2895 |
| Qy | 1696 CAGGAGAGAACCAACTCTGAGTGGCGATATGACACAGGGAGATTACTGTG | 1755 | Db | 2883 GCTTCACCAAACTCTACAGAGTCAGTAGTCAGTCAGTCAGTCAG | 2942 |
| Db | 1743 AATGAACCTAAACTCACTGAGCTGCCACGGGATGATCTCAGGAGCTTG | 1802 | Qy | 2896 ACCAGAGGGACTCTCTGTCGGCTCAGCCCTTCTCCAGACATCAGCTCCCTGTC | 2955 |
| Qy | 1756 AGTTTGTGATGAACTGAGCTGAGGCGCTAACAGACAGGGAGTACTCAGTG | 1815 | Db | 2943 ACCGGAGGATTCTCTGTCAGCCCTGGGAGATCTCCATGAGCTGCAAGTGGAG | 3002 |
| Db | 1803 ACCTCTGCTGAGAACTGAGCTGAGGAGCTGAGCTGAGTACGATCAG | 1862 | Qy | 3016 AAATCCACTCCAGCTGAGGCTTGGACCAATATGCTGACCTTGGAG | 3075 |
| Qy | 1816 GAGAGAGAACCAACTCTGAGTGGCGATATGACACAGGGAGATTACTGTG | 1875 | Db | 3063 AAGTTCACACTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAG | 3122 |
| Db | 1863 GAGAAGACTGAGAAACTGAGTGGCGCTAACAGACAGGGAGTACTCAGTG | 1922 | Qy | 3076 AAAACAGACCTGGACCTGAGAGGAGATGCCCCAGACTGCACTCACCTGTC | 3135 |
| Qy | 1876 GAGGAGAGAACCAACTCTGAGTGGCGATATGACACAGGGAGATTACTGTG | 1935 | Db | 3123 AAATGCTACTTGGAAATCAGATACATCTCTTGTCCCCAGACTGCACTCTCTAAGCCAG | 3182 |
| Db | 1923 GTCCTCCAGCCATGTTGAAATGTCATCCACCGACCTCTGGAGCTGCGCAG | 1982 | Qy | 3136 CCCTCTGGAGACCCCTGGAGATGAGACCTACACCTACGACCTTCAGGCAACCT | 3195 |
| Qy | 1936 ACTCTTCTGCTCTGACTTTCACTTCAGCCAGCTCACTCCGGAGGCC | 1995 | Db | 3183 CCTTCCTGGAGACCTCTCCGGACAGACTACACCTACGACCTTCAGGCAACCT | 3242 |
| Db | 1983 ACTCTTCTGCTCTGCCCTGACTTTCACTTCAGTCAGACTCACGCTCAGCC | 2042 | Qy | 3196 AACTGGGCACTGACTTCCTGAGACCTACACCTACGACCTTCAGGCAACCT | 3255 |
| Qy | 1996 AGTGTCTCTGAGAACATAATGCTGGAGCTGTCGGCCAGAAAGACGACTGTCATAAGTT | 2055 | Db | 3243 AATGGGGAGCTGACTCTGTCAGTCAGTCAGTCAGTCAGTCAGTCAG | 3302 |
| Db | 2043 AGCCCTCTGAGAACATAATGCTGGAGCTGTCGGCCAGAAAGACGACTTCAGGAAATTT | 2102 | Qy | 3256 GTCCACGGCTCCGACCGAGACATCAAGTGTGGCTCTCTGGGG | 3315 |
| Qy | 2056 GAAACAGAGATATCACAGTCGAGCCATGGCCCTTCAGGAGCTCAAG | 2115 | Db | 3303 GTTCAGGACTCCGACCATCGACATCAAGTGTGGCTCTCTGGGG | 3362 |
| Db | 2103 GAATCAAGGTCTCATCATGTCCTCAACACAGACCTCACCTTCTGGGG | 2162 | Qy | 3316 ACAGACTGGAGCTGAGCAACAAACTCCAGTACCCACATCTGAGGGACTC | 3375 |
| Qy | 2116 AACAGCATGAGGTCTGGGAGCTGGCCATGGCCCTTCAGGAGCTCAAG | 2175 | Db | 3363 AGCCACGGACTCCGACCAAGTGTGGCTCTCTGGGG | 3422 |
| Db | 2163 AACAGACACTCTGGAGACATGAGACTCTGAGGAGAACCTCCCGCTCTGCTCA | 2222 | Qy | 3376 TCTTGTGAGATGAGGGACTCTGTCAGACAGACTGCTGCTGCTG | 3435 |
| Qy | 2236 GATTCCTGACCGCTGGCAATGAGCTGAGGAGACCTCCCGCTCTGCTCA | 2295 | | | |
| Db | 2283 GACTCTGACTCTGGCAATGAGACTGAGGAGATCTGCTGCTGCTG | 2342 | | | |
| Qy | 2296 ACACAGTATGGGGACTCTGTCAGTCAGGAGGGAGGGCTCCCTGAGAATGTCAC | 2355 | | | |

Db 3423 TCTGGAGCATGGAGGATGGAACGTGGAGACCACTACGCCAATCTG 3482 DT 16-DEC-2002 (first entry)
 Qy 3436 AAGAGTCACCCCTTACTCTCTGGCTCTTACCAACACTCGGGAGAGGA 3495 XX DE Human phospholipase protein encoding cDNA SEQ ID NO:1.
 Db 3483 AGAGAGTTCAACCTTCCATCTCTGGATTCACCGTAGGAGACAGGA 3542 XX Human; phospholipase; enzyme; chromosome 2; gene; ss.
 Qy 3496 CTAATGGCAGGGAGGGCCAGAGCTAGGACATCCACGCCAGCTGGACCTG 3555 XX OS Homo sapiens.
 Db 3543 TTAATGGCAGAGAAACGCCATCACATCACTACAGGAGACTGGACCTG 3602 XX FH Key Location/Qualifiers
 Qy 3556 GTAGAGGGATGAAACAGCCGAGACATCACTGGAGAAGACTGGTACA 3615 FT 5'UTR 1.175 /*tag= a
 Db 3603 GTGAGAGATGAAACGCCATCACATCACTACAGGAGACTGGACCTG 3662 FT CDS 176..1594 /*tag= b
 Qy 3616 CTCTCTCATGGAGAACAGACTCTGTCTTACTGTGAGAAGCTGGTACA 3675 FT /product= "phospholipase"
 Db 3663 CTCTCTCATGGAGAACAGACTCTGTCTTACTGTGAGAAGCTACACC 3722 FT /*tag= c
 Qy 3676 ACCGAATATGTCAGCACATCCACAGGCCCTGGACATCCTCTGGAGAGCTCCAGG 3735 PN w020262977-A2.
 Db 3723 AGGGAGATGTCACTGGAGACTCTGTGAGAAGCTGGAGGCTTGTGAGG 3782 XX PD 15-AUG-2002.
 Qy 3736 GCTTGTCAGAACGCTGGAGACTCTGTGAGAATGGAGCTAGCTGGCT 3795 XX PF 28-JAN-2002; 2002W0-US02302.
 Db 3783 GTTTCATCAAGGGTGGAGATGGAGCTGGAGCTGGCTTGTGAGG 3795 XX PR 08-FEB-2001; 2001US-0778961.
 Qy 3796 AAATGTGCATGC--TGGCAGCTCAGAACACTGCACTGCTCAGAC 3852 PA (PEKE) PE CORP NY.
 Db 3843 AAATGTGCATGCCTGGCTGGCTGGCTGGCTTAACGCTCATTAAAC 3902 PI Yan C, Ketchum KA, Di Francesco V, Beasley EM;
 Qy 3853 TCCTGGAGAACAGAACTGAACTGGAACTTCCAGATGGCTCCAGT 3912 DR DR WI; 2002-682698/73.
 Db 3903 CTCTGGCATGGAGGAACTGAACTGGAACTTCCAGTCCTGG 3962 XX P-PSDB, ABP53556.
 Qy 3913 TCTCTCTACTGCAACAAATACACAAAGCTGAGGACTTGGGTGTTGGAG 3972 XX New human phospholipase proteins, useful for the development of human tissue
 Db 3963 CTCTGGACTGGCTGGTACACTGGCTGGCTGGCTGGCTGG 4022 PT therapeutics and diagnostic compositions, drug screening assays, tissue
 Qy 3973 TTCCAAACACACTACCCACTGAACTGGAGAGGGACACTGCCTCTCC 4032 PT typing and pharmacogenomic analysis -
 Db 4023 TTCTGGAAATACCTTGTCCCATCTGGATGAGCTGGGGCTCGACTCTCT 4082 XX
 Qy 4033 GAGGACTGTTTACTCTCAAGACGGAGGGACATCCGAGATGCTGAC 4092 CC Claim 4; FIG 1A; 95pp; English.
 Db 4083 GAGGACTGTTTCCACTCTCACTGGATGAGCTGGGGCTCGACTCTCT 4142 CC The present sequence encodes a human phospholipase protein (I) located on
 Qy 4093 AACCTGCTGAGGACTGGCCGGAGACTTCTCAACAACTCACCACGGAGCC 4152 CC chromosome 2. (I) can be used for identifying agents that modulate its
 Db 4143 AACATGCTGAGGACTGGAGAGAACCTCCAACTACCTACATACGGAGCC 4202 CC function or activity where the agent is useful for treating a disease
 Qy 4153 AACCTCAACTGCCCTCTCTGAGAGGCCCTACTCTACCCCTGGGAGACCGATG 4212 CC or condition mediated by a the human phospholipase protein. (I) peptides
 Db 4203 AACCTCAAGTGTCCCTCGCTGAAACCTTCTCTCAACTGTGGAGACCTCAGATT 4262 CC can be used in substantial and specific assays related to functional
 Qy 4213 CTCCAGGAGGCTGAGGAGAACCTCCAACTACCTACATACGGAGCC 4272 CC information or the peptide sequences, to raise antibodies or to elicit
 Db 4263 CTCTAGACAAAGCTGAGGAGAACCTCAAATACCTACTGGCACTGGCTCA 4322 CC immune response, as reagents in assays that determine the levels of
 Qy 4273 GAGGCTGGCTGTGGAGGACATGGGAGGAGGGTGGAGGAGGGTGC 4332 CC protein in biological fluids, and as markers for tissues where the
 Db 4323 GAGGCTGGCTGTGGAGGAGGGTGGAGGAGGGTGC 4382 CC corresponding protein is expressed. Nucleotide sequences encoding (I)
 Qy 4333 CGGAGGGAGA 4343 CC can be used as probes, primers and chemical intermediates in biological
 Db 4383 CAATAGAAGGA 4393 CC assays for constructing recombinant vectors and expressing antigenic
 Qy Sequence 1835 BP; 458 A; 549 C; 474 G; 354 T; 0 other; CC portions of the protein. (I) and nucleic acid molecules encoding it can
 Db Matches 1582; Conservative 0; Mismatches 0; Indels 190; Gaps 3; CC be used in the identification of therapeutic proteins and may serve as
 Qy 2618 CTGCAGCCACTTGTGACCAATCTCCGAATGCCTGGACCTGAGAGCTT 2671 CC models or targets for the development of human therapeutic agents that
 Db 1 CTGAGCCACTTGTGACCAATCTCCGAATGCCTGGACCTGAGAGCTT 60 CC modulate phospholipase activity in cells and tissues that express the
 Qy 2726 GGAGGTTCTGGAAACCCAGACAGTGCCTGGACAGGGCTTGTGA 2785 CC phospholipase, such as in kidney, blood, lung, brain glioblastoma,
 Db 121 GCGAGGTTCTGGAAACCCAGACAGTGCCTGGACAGGGCTTGTGA 168 CC prostate, colon or leukocytes.
 Qy Sequence 1835 BP; 458 A; 549 C; 474 G; 354 T; 0 other;
 Db Best Local Similarity 31.1%; Score 1362; DB 24; Length 1835;
 Qy Matches 1582; Conservative 0; Mismatches 0; Indels 190; Gaps 3;
 Db 2618 CTGCAGCCACTTGTGACCAATCTCCGAATGCCTGGACCTGAGAGCTT 2671
 Qy 1 CTGAGCCACTTGTGACCAATCTCCGAATGCCTGGACCTGAGAGCTT 60
 Qy 2726 GGAGGTTCTGGAAACCCAGACAGTGCCTGGACAGGGCTTGTGA 2785
 Db 121 GCGAGGTTCTGGAAACCCAGACAGTGCCTGGACAGGGCTTGTGA 168

RESULT 9
 ABQ82234 ID ABQ82234 standard; cDNA; 1835 BP.
 XX AC ABQ82234;
 XX

| | | | | | | | |
|----|------|--|------|-----------|------------------------------|--|------|
| QY | 2786 | ACTGGGTTCGACCCAGAGCTCCAAAGAGCTAGCCAGGCTGGGGCTTCAGCC | 2845 | QY | 3866 | AAGACTCTAGAAGAGTCACTGAACTCCACATGCCATCCTCCAGTTCTCTACTCGGC | 3925 |
| Db | 169 | - | - | Db | 1083 | AAGAACCTGAGAAAGTGAACTGCACTCCAGATGCCATCAGTTCTCTACTCGGC | 1142 |
| QY | 2846 | GACCTACGGGACAGCATGGAGCTGGTGGGGCTAGGGCTATGACAGCAGGAGG | 2905 | QY | 3926 | ACCAATACACAGCGCTGGAGCTTGCGGTGGGGCTGGGGCTGGGGCTTCAGACACAC | 3985 |
| Db | 169 | - | - | Db | 1143 | ACCAATACACAGCGCTGGAGCTTGCGGTGGGGCTGGGGCTGGGGCTTCAGACACAC | 1202 |
| QY | 2906 | ACCTCTCTGGGGCTGAGCCTTCAGAACATCGAGTCCCCTCTGGGGATG | 2965 | QY | 3986 | TCAACCCACTGAGCAGGAGGGACACTGAACTTCAGCTTCAGGAGACTGTTTC | 4045 |
| Db | 219 | ACCTCTCTGGGGCTGAGCCTTCAGAACATCGAGTCCCCTCTGGGGATG | 273 | Db | 1203 | TCAACCCACTGAGCAGGAGGGACACTGAACTTCAGCTTCAGGAGACTGTTTC | 1262 |
| Db | 274 | - | - | QY | 4046 | ACTTCAGACGGGGCATGGGGCATCGACTCTGGGGCTTCAGCTTCAGGAGAC | 4105 |
| QY | 3086 | CCTGGGACTGAGCAGAGATGCCATCACCTGTCACACTAGAACTAGCCCTCTGA | 3025 | Db | 1263 | ACTTCAGACGGGGCATGGGGCATCGACTCTGGGGCTTCAGCTTCAGGAGACTGTTTC | 1322 |
| Db | 303 | CCTGGGACTGAGCAGAGATGCCATCACCTGTCACACTAGAACTAGCCCTCTGA | 3145 | QY | 4106 | CAGTGCCCCGAGACTACTCCACAACTACCCAGAGCGGAGAACATCAAGTGCC | 4165 |
| QY | 3146 | GAACCCCTGGGATAGTAACATACAGTACCCATCAAGGCCATGAGAACTGGGCA | 3205 | Db | 1323 | CTGTGGCCGAGACTACCTCCACACTTACCCAGGAGGAGGAGCTGAGGAGGAC | 1382 |
| Db | 363 | GAACCCCTGGGATAGTAACATACAGTACCCATCAAGGCCATGAGAACTGGGCA | 422 | QY | 4166 | CCTCTCTGAGAGCCCTTACCTCTACACCTCGGAAAGAGCGGATGTCAGGAG | 4225 |
| QY | 3206 | GTGACTTCTGTGTTACAGAGTGGAGGTTCCATACTGTTCAACTCTGTCACCG | 3265 | Db | 1383 | CTCTCTGAGAGCCCTTACCTCTACACCTCGGAAAGAGCGGATGTCAGGAG | 1442 |
| Db | 423 | GTGACTTCTGTGTTACAGAGTGGAGGTTCCATACTGTTCAACCTGTGACCCAGC | 482 | QY | 4226 | CTGAGAGGCCCCGAGGTGCTACTGGGTGTCAGGGAGGTTGGGAGGAGTC | 4285 |
| Qy | 3266 | TGGACCGAGACATCAAAGGGGGCTGGGGACTCTGACTAGGAGCTGG | 3325 | Db | 1443 | CTGAGAGGCCCCGAGGTGCTACTGGGTGTCAGGGAGGTTGGGAGGAGTC | 1502 |
| Db | 483 | TGGACCGAGACATCAAAGGGGGCTGGGGACTCTGACTAGGAGCTGG | 542 | QY | 4286 | TGGGGCATCATGGGAGAGTGGGTCTGGAGGGTGGGGAGGAGTC | 4345 |
| QY | 3326 | GAACCTGACCAACAACCTCCAGGACCTTACCCACATCTGGGGACTCTGGAGCA | 3385 | Db | 1503 | TGGGGCATCATGGGAGAGTGGGTCTGGAGGGTGGGGAGGAGTC | 1562 |
| Db | 543 | GAACCTGACCAACAACCTCCAGGACCTTACCCACATCTGGGGACTCTGGAGCA | 602 | QY | 4346 | CTCCAATGAGCTGCCACTCTGGGCTCTAG | 4377 |
| QY | 3386 | TGGAGGGGATGGAACTTGGGACTCACCAACACTGGCCACATCTGGAGAGTC | 3445 | Db | 1563 | CTCCAATGAGCTGCCACTCTGGGCTCTAG | 1594 |
| Db | 603 | TGGAGGGGATGGAACTTGGGACTCACCAACACTGGCCACATCTGGAGAGTC | 662 | RESULT 10 | | | |
| QY | 3446 | ACCTTACCTCTGGCTCTTACAGGACCTGGGACGGGAGGAGCACTAATGG | 3505 | ID | ABX97049 | standard; cDNA; 1624 BP. | |
| Db | 663 | ACCTTACCTCTGGCTCTTACAGGACCTGGGACGGGAGGAGCACTAATGG | 722 | XX | ABX97049; | | |
| QY | 3506 | CGCGGAGGGGCGAGCTAGGACATGCCAGGCCCTGGGAGCTGGAGGAA | 3565 | DT | 20-MAY-2003 | (first entry) | |
| Db | 723 | CGCGGAGGGGCGAGCTAGGACATGCCAGGCCCTGGGAGCTGGAGGAA | 782 | DE | Human | NON24b cDNA. | |
| QY | 3566 | TGAAACAGGCCAGATCAACTGGAGAAGACTGAGTGGTACACTCTATG | 3625 | XX | NOMX: | cystostatic; cardiotonic; antiarrhoesclerotic; antiasthmatic; cancer; hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine; human; gene: ss. | |
| Db | 783 | TGAAACAGGCCAGATCAACTGGAGAAGACTGAGTGGTACACTCTATG | 842 | OS | Homo sapiens. | | |
| QY | 3626 | GGTCAACAGACTGTGCAATTCTGTGAGAATCCGGAGCCACTTGCCAGGAAATG | 3685 | XX | W0200272757-A2. | | |
| Db | 843 | GGTCAACAGACTGTGCAATTCTGTGAGAATCCGGAGCCACTTGCCAGGAAATG | 902 | XX | 19-SEP-2002. | | |
| QY | 3686 | TTCAGCACATCCACAGGCCCTGGACATCTCTGGAGGCTCCAGGGGTTCTCA | 3745 | PD | | | |
| Db | 903 | TTCAGCACATCCACAGGCCCTGGACATCTCTGGAGGCTCCAGGGGTTCTCA | 962 | PP | 08-MAR-2002; 2002WO-US06908. | | |
| QY | 3746 | ACGTGGTGGAGGCTCATGGCTGGCTGGCTGGAGCTGGAGCTGGAGCTGGAGC | 3805 | PR | 08-MAR-2001; 2001US-274101P. | | |
| Db | 963 | ACGTGGTGGAGGCTCATGGCTGGCTGGAGCTGGAGCTGGAGCTGGAGC | 1022 | PR | 08-MAR-2001; 2001US-274201P. | | |
| QY | 3806 | TCTGGGAGCTAGAACACTGACTCTGGCTAGACACTGAGGAAAGTGGAGAGC | 3865 | PR | 08-MAR-2001; 2001US-274322P. | | |
| Db | 1023 | TCTGGGAGCTAGAACACTGACTCTGGCTAGACACTGAGGAAAGTGGAGAGC | 1082 | PR | 09-MAR-2001; 2001US-27489P. | | |
| PR | | | | PR | 12-MAR-2001; 2001US-27525P. | | |
| PR | | | | PR | 13-MAR-2001; 2001US-27557P. | | |
| PR | | | | PR | 13-MAR-2001; 2001US-27559P. | | |
| PR | | | | PR | 13-MAR-2001; 2001US-275601P. | | |
| PR | | | | PR | 14-MAR-2001; 2001US-276000P. | | |
| PR | | | | PR | 16-MAR-2001; 2001US-27676P. | | |
| PR | | | | PR | 19-MAR-2001; 2001US-27694P. | | |
| PR | | | | PR | 20-MAR-2001; 2001US-277321P. | | |

PR 20-MAR-2001; 2001US-277327P.
 PR 21-MAR-2001; 2001US-277791P.
 PR 22-MAR-2001; 2001US-277833P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 26-MAR-2001; 2001US-278894P.
 PR 27-MAR-2001; 2001US-278999P.
 PR 27-MAR-2001; 2001US-279036P.
 PR 28-MAR-2001; 2001US-279344P.
 PR 30-MAR-2001; 2001US-27995P.
 PR 30-MAR-2001; 2001US-280233P.
 PR 02-APR-2001; 2001US-280802P.
 PR 02-APR-2001; 2001US-280822P.
 PR 02-APR-2001; 2001US-280900P.
 PR 04-APR-2001; 2001US-281194P.
 PR 13-APR-2001; 2001US-283675P.
 PR 30-APR-2001; 2001US-287424P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 03-MAY-2001; 2001US-288528P.
 PR 15-MAY-2001; 2001US-291190P.
 PR 16-MAY-2001; 2001US-291099P.
 PR 16-MAY-2001; 2001US-291240P.
 PR 30-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294889P.
 PR 31-MAY-2001; 2001US-294899P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 19-JUN-2001; 2001US-299310P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 11-JUL-2001; 2001US-304355P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 16-AUG-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-335301P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 14-NOV-2001; 2001US-333272P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-0092900.
 XX PA (CURA-) CURAGEN CORP.
 XX PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CBA, Li L;
 PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kerkuda R;
 PI Ratturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
 PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Leite M, Zhong H;
 PI Anderson D, Spaderna SK, Catterson B, Burgess C, Leite M;
 PI Alsobrook JP, Lepley DM, Rieger DK;
 XX DR WPT; 2002-723332/78.
 DR -PSDB; ABUS082.
 XX PT NOX polypeptides and polynucleotides, useful for preventing or
 PT treating a disorder associated with aberrant NOX expression or
 PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
 PT bronchial asthma -
 PS Claim 13; Page 169; 1103pp; English.

CC preventing or treating a disorder associated with aberrant NOX expression or activity e.g. cancer, hypertension, atherosclerosis, cardiomopathy or bronchial asthma. The products of the invention can be used for gene therapy or in a vaccine. ABX9708-ABX9715 are cDNA fragments amplified and isolated by the PCR primers and probes represented in ABX1460-ABX1342 and ABX3718-ABX3753. ABX7008-ABX97185 encode the NOX proteins described in ABU65041-ABU65218.

Sequence 1624 BP; 422 A; 421 C; 401 G; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

SP Sequence 1624 BP; 422 A; 401 C; 380 T; 0 other;

CC preventing or treating a disorder associated with aberrant NOX expression or activity e.g. cancer, hypertension, atherosclerosis, cardiomopathy or bronchial asthma. The products of the invention can be used for gene therapy or in a vaccine. ABX9708-ABX9715 are cDNA fragments amplified and isolated by the PCR primers and probes represented in ABX1460-ABX1342 and ABX3718-ABX3753. ABX7008-ABX97185 encode the NOX proteins described in ABU65041-ABU65218.

Query Match Similarity 26.5%; Score 1160.6; DB 24; Length 1624; Matches 1163; Conservative 99.7%; Pred. No. 0; Mismatches 4; Indels 0; Gaps 0;

QY 1161 CGCGGCTGACATCAACGGTATTGGGACTGGCTTCAGTCAGCAGGCCCTGGCGAG 1220

Do 61 CGGGTCCACACATGGAATTGGAGCTCTGACTAGTGGAGCCCTGGCGAG 120

QY 1281 C37GGGGAGAGTGGAGAATGGGACAGGGTACCTGGAGAGACTGGAGATAATAC 1340

Do 121 CGTGAGGAGATGAGAACGAGCTTACCCGGACAGCTTCTGGGAACTCCGGGAACTTCTGGGAAATT 180

QY 1341 CAACCTTCTCCAGAAGGGCTTCCTGTTGGACTGGAGAAGAACCTGGCTTAATGCCCT 1400

Do 241 CTAAACAGGCTGTGGAGGGCCGAGCTGAGGAACTACCTGTCGCCGCGAGGCT 300

QY 1461 GGTGGACCTGAGAATGAGACGAGGATACCTTGGAGAGACTGGAGATAATAC 1520

Do 301 GGTTGACCTGAGAATGAGACGAGGATACCTTGGAGAGCTGGAGATAATAC 360

QY 1521 CCTGTTTATGGGGCATAGACCTCTGTTGAGTTCATGATGTCATGTCCTCCACTATCTCC 1580

Do 421 CCAGAACCTCACAGACACATGGAAAGGCCCTGGACATCCCTGCATGTCAGGTTCCCG 480

QY 1641 GGGATTGTAACCTGGTACGGTGTGAGTCTGAGTCACTGGAGGACTTACCCGA 1700

Do 481 GCGATTGTAACCTGGTGTGAGTCACTGGAGGACTTACCCGA 540

QY 1701 GAAAAGCTACTGCCAGGATGTCCTGAGTCCTGTCCTGTCCTGTC 1760

Do 541 GAAAAAGCTACTGCCAGGATGTCCTGAGTCCTGTCCTGTCCTGTC 600

QY 1761 TGATGAACTCACAGACAGCTTGCTACCCATGAACTTCAGGAGTTTCAGGAGA 1820

Do 601 TGATGAACTCACAGACAGCTTGCTACCCATGAACTTCAGGAGTTTCAGGAGA 660

QY 1821 GACCCACCACTGAGAGCTGGGAGATGACAGCAAGGGAGATTACTGGGTT 1880

Do 661 GACCCACCACTGAGAGCTGGGAGATGACAGCAAGGGAGATTACTGGGTT 720

QY 1881 GCAGCGCTCTTGAACAGCTGGACATGCCAAAGACCTGGAGAGTTTCAGGAGA 1940

Do 781 TTCTTCCTCTGAGCTTGCTACCTGGAGGACTTCAGTCAGGAGCCAGTC 840

QY 721 GCAGCGCTCTTGAACAGCTGGACATGCCAAAGACCTGGAGGAGTTTCAGGAGA 780

QY 1941 TTCTTCCTCTGAGCTTGCTACCTGGAGGACTTCAGTCAGGAGCCAGTC 2000

Do 841 TCTCTGGACAATATGGAGGCTGTGGAGAGAGGACTGGCTCATAGTTGAAA 900

| | | | |
|----|--|---|---|
| QY | 3537 AGCCCAGGCTGGGAACTGTAGGGAAACGCCCACTACACCTGGGAAA 3596 | CC | antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic; |
| Db | 614 AGCCCCGGCTGGGAACTGTAGGGAAACGCCCACTACACCTGGGAAA 673 | CC | antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; |
| QY | 3597 AGACTGGAAAGCTGGTCAACTCTCATGGGTCAAGACTGTGTCAATCTGGAAA 3656 | CC | anticapinsonian and immunomimetic. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorder, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and neurological disorders. |
| Db | 734 TCCGGAGGCCACTTGCCAGGAATATGTTAGCAGCACATCCACAGGCCCTGGACATCT 793 | CC | Sequence 802 BP; 206 A; 236 C; 211 G; 149 T; 0 other; |
| | | XX | Query Match 15.2%; Score 667.4; DB 22; Length 802; Best Local Similarity 99.9%; Pred. No. 3_5e-177; Matches 668; Conservative 0; Mismatches 1; Indels 0; Gaps 0; Matches 668; Conservative 0; Mismatches 1; Indels 0; Gaps 0; |
| AC | AHH99765; | XX | 3057 GCTTGAAACACTTGGAACGAAACAGAGACCTTGAGAGCAGATGCCATAC 3116 |
| AC | AHH99765; | XX | Db 134 GCTTGAAACACTTGGAACGAAACAGAGACCTTGAGAGCAGATGCCATAC 193 |
| DE | Human protein encoding cDNA sequence SEQ ID NO:600. | XX | QY 3117 CTGTCCTCACTCAGAATGGGCCCTTCCTGAGAACCCCTGGAAACTGACTACAGTRACC 3176 |
| XX | Human; cancer; ulcer; HIV infection; human immunodeficiency virus; antiinflammatory; antineoplastic; vulnerary; antiallergic; osteopathic; antibiotic; endocrine; cardiot; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianemic; anaemia; antiaggregant; haemostatic; vulnerary; antiallergic; osteopathic; eczema; dermatological; antiidiotic; cyrostatic; neuroprotective; antidepressant; notropic; antiparkinsonian; infection; immunosimulant; gene therapy; antisense therapy; vaccine; inflammation; antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis; cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity; genetic disease; haemopoietic disorder; platelet disorder; asthma; thrombocytopaenia; osteoporosis; severe combined immunodeficiency; allergic rhinitis; diabetes; multiple sclerosis; depression; disorder; Alzheimer's disease; Parkinson's disease; neurodegenerative disorder; neurological disorder; ss. | Db 194 CTGTCCTCACTCAGAATGGGCCCTTCCTGAGAACCCCTGGAAACTGACTACAGTRACC 253 | |
| XX | Homo sapiens. | XX | QY 3177 CATCAAGCAGCATTGGAACACTGGGGAGTACTCTGTGAGCTAGNGGGAGCTC 3236 |
| OS | W00015345 A2. | XX | Db 3237 CAATAGTGTCCACCTCTGTGTCACCAAGCTCGACCGACGAGCATCAAGTGTGGCC 3296 |
| OS | | XX | Db 314 CATTAGTGTCCACCTCTGTGTCACCAAGCTCGACCGACGAGCATCAAGTGTGGCC 373 |
| XX | | XX | QY 3227 CTCTGGTACTCTGTGACTACAGCAGTGGAACACTCAGTGACCTTAC 3356 |
| XX | | XX | Db 374 CTCTGGTACTCTGTGACTACAGCAGTGGAACACTCAGTGACCTTAC 433 |
| PN | | XX | QY 3357 CACATCTGGAGGGACTCTCTGGACCATGGACAAACAACCTCGACGTTAC 3416 |
| PN | | XX | Db 434 CACATCTGGAGGGACTCTCTGGACCATGGACAAACAACCTCGACGTTAC 493 |
| PD | 26-JUL-2001. | XX | QY 3417 CACACTGCCAACATCTGGAGAGTCAACCTTACCCCTTGCTTCTAACAGAC 3476 |
| XX | | XX | Db 494 CACACTGCCAACATCTGGAGAGTCAACCTTACCCCTTGCTTCTAACAGAC 553 |
| PF | 22-DEC-2000; 2000WO-US35017. | XX | QY 3477 CTGGAGGGAGCAGCAGACTAAATGTCAGGGACGGGACATCCAGTC 3536 |
| PR | 23-DEC-1999; 9905-0471275. | XX | Db 554 CTGGAGGGAGCAGCAGACTAAATGTCAGGGACGGGACATCCAGTC 613 |
| PR | 21-JAN-2000; 2000US-048725. | XX | QY 3537 AGCCCAAGCTGGGACTCTGGTAGAGGAATGAAACAGCCCGACATCACTGGGAA 3596 |
| PR | 25-APR-2000; 2000US-052317. | XX | Db 614 AGCCCAAGCTGGGACTCTGGTAGAGGAATGAAACAGCCCGACATCACTGGGAA 673 |
| PA | (HYSE-) HYSEQ INC. | XX | QY 3597 AGACTGGAAAGCTGGTCAACTCTCATGGGTCAACGACTGTGTCAATCTGTGAGAA 3656 |
| PA | Tang YT, Liu C, Drmanac RT; | XX | Db 674 AGACTGGAAAGCTGGTCAACTCTCATGGGTCAACGACTGTGTCAATCTGTGAGAA 733 |
| PI | | XX | QY 3657 TCCGGAGGCCACTTGCCAGGAATATGTTAGCAGCATCCACAGGCCCTGGACATCT 3716 |
| XX | | XX | Db 734 TCCGGAGGCCACTTGCCAGGAATATGTTAGCAGCATCCACAGGCCCTGGACATCT 793 |
| PS | Claim 1; Page 634; 1217PP; English. | XX | QY 3717 CTCTGAGGA 3725 |
| XX | | XX | Db 794 CTCTGAGGA 802 |
| CC | AHH9916 to AHH9904 encode the human proteins given in AM22225 to AM25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antiinflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiot; | XX | RESULT 13 |
| CC | central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary; | XX | ADD37412 |

| | | | |
|-----------------------|---|--|--|
| ID | AD37412 | standard; DNA; 608 BP. | 13.2%; Score 576.2; DB 24; Length 608; |
| XX | | Best Local Similarity 97.8%; Fred. No. 1.15e-151; | |
| AC | AAD37412; | Matches 584; Conservative 0; Mismatches 13; Index 0; Gaps 0; | |
| XX | | | |
| DT | 27-AUG-2002 (first entry) | | |
| XX | | | |
| DE | Human phospholipase-like enzyme encoding DNA #2. | | |
| XX | | | |
| KW | Human; phospholipase-like enzyme; cancer; inflammation; Pick's disease; cardiovascular disorder; central nervous system disorder; brain injury; chronic obstructive pulmonary disease; cerebrovascular disease; dementia; Alzheimer's disease; Parkinson's disease; corticobasal degeneration; motor neuron disease; Huntington's disease; Creutzfeld Jacob dementia; schizophrenia; Korsakoff's psychosis; pain; epilepsy; multiple sclerosis; sciatica; stroke; age associated memory impairment; allergy; asthma; allergic rhinitis; hay fever; atopic dermatitis; cardiovascular disease; anaphylaxis; inflammation; acute respiratory distress syndrome; diabetes; chronic obstructive pulmonary disease; emphysema; obesity; anorexia; overweight; cachexia; bulimia; hypertension; coronary artery disease; type-II diabetes; hyperlipidaemia; gall bladder disease; osteoarthritis; gout; sleep apnoea; respiratory problem; polycystic ovarian syndrome; thrombolytic disease; reduced fertility; pregnancy; stress incontinence; hirsutism; menstrual irregularity; depression; enzyme; ds. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |
| PN | WO200231161-A2. | | |
| XX | | | |
| PD | 18-APR-2002. | | |
| XX | | | |
| PP | 09-OCT-2001; 2001WO-EP11641. | | |
| XX | | | |
| PR | 10-OCT-2000; 2000US-238445P. | | |
| PR | 26-DEC-2000; 2000US-257293P. | | |
| XX | | | |
| PA | (FARB) BAYER AG. | | |
| XX | | | |
| PI | Zhu Z; | | |
| XX | | | |
| DR | WPI; 2002-426287/45. | | |
| XX | | | |
| PT | New human phospholipase-like enzyme polypeptide useful for screening agents, and in the treatment of cancer, inflammation, diabetes, obesity, a central nervous system disorder, or a cardiovascular disorder - | | |
| PT | | | |
| PS | Disclosure; Fig 5; 144pp; English. | | |
| XX | | | |
| CC | The present invention relates to novel human phospholipase-like enzymes and polynucleotides encoding such proteins. Sequences of the invention are useful for producing a medicament for modulating the activity of phospholipase in a disease such as cancer, inflammation, cardiovascular disorders, chronic obstructive pulmonary diseases, central nervous system (CNS) disorders such as brain injuries, cerebrovascular disease, dementia (Alzheimer's disease), Parkinson's disease, corticobasal degeneration, motor neuron disease, Pick's disease, Huntington's disease, Creutzfeld Jacob dementia, schizophrenia with dementia, Korsakoff's psychosis, pain associated with CNS (e.g. epilepsy, failed back surgery syndrome, sciatica), multiple sclerosis, stroke, age associated memory impairment, allergic disease including asthma, allergic rhinitis (hay fever, atopic dermatitis, anaphylaxis and inflammation, cardiovascular disease, chronic obstructive pulmonary disease, acute respiratory distress syndrome, gout, diabetes, emphysema or obesity). They are also used for treating anorexia, overweight, cachexia, bulimia, hypertension, type-II diabetes, coronary artery disease, hyperlipidaemia, gall bladder disease, osteoarthritis, sleep apnoea and respiratory problems, cancer (e.g. breast, prostate, colon cancer), thrombolytic disease, reduced fertility, polycystic ovarian syndrome, complications of pregnancy, menstrual irregularity, hirsutism, stress incontinence and depression. The present sequence is human phospholipase like enzyme encoding DNA. | | |
| SQ | Sequence 608 BP; 140 A; 187 C; 158 G; 123 T; 0 other; | | |
| Query Match | 13.2%; Score 576.2; DB 24; Length 608; | | |
| Best Local Similarity | 97.8%; Fred. No. 1.15e-151; | | |
| Matches | 584; Conservative 0; Mismatches 13; Index 0; Gaps 0; | | |
| Db | 2654 TGGACGTCCTGCATAGAGGGCTGCCAGAGCTCTGCTCAACTCTGGAGCT 2713 | 11 TTGATGGCGCATAGAGGGTGCCTGCAGAGCTCTGCTCAACTCTGGAGCT 70 | |
| Qy | 2714 CCACTATCATCGCGAGGTCTTCGGGAAACCGAGACAGTGCCAGTSGAGAGCCA | 2774 CGGTTTGTGTACTGGCTTCGGGAGACTCCAGCTACGGCTGG 2833 | |
| Db | 71 CCACATCATCGCGAGGTCTTCGGGAAACCGAGACAGTGCCAGTSGAGAGCCA | 130 131 CGGTTTGTGTACTGGCTTCGGGAGACTCCAGCTACGGCTGG 190 | |
| Qy | 2834 AGGCCTTCAGCGAGCTAACCGAGCTACGGAGCACAGTGCCAGTSGAGAGCCA | 2894 ACACCGAGGAGACTCTCTGCTGGCTGCAGCCAGTCAGCCAGTG 2953 | |
| Db | 191 AGGCCTTCAGCGAGCTAACCGAGCTACGGAGCACAGTGCCAGTSGAGAGCCA | 250 251 ACACCGAGGAGACTCTCTGCTGGCTGCAGCCAGTCAGCCAGTG 310 | |
| Qy | 2954 TCTTGGCGATGGCTCCAGATGCTCTTCTTGGCCAGCTACGGCTGG 3013 | 371 AGAAATTCCAGCTCCAGCTGGCCAGGCCATTCAGCTGAACCAATATGCTGAACCAACT 3073 | |
| Db | 311 TCTTGGCGATGGCTCCAGATGCTCTTCTTGGCCAGCTACGGCTGG 370 | 3074 GCACAAACAGAGACCTGGACCTGAGAGATGCCATCACCTGGTCCCTCTGCCCCAGACTGCTACATCCACCAAATC 430 | |
| Qy | 3014 AGAAATTCCAGCTCCAGCTGGCCAGGCCATTCAGCTGAACCAATATGCTGAACCAACT 3073 | 431 GCACAAACAGAGACCCCTGGAGACAGAGATGCCATCACCTGGACATCCACACTGAGATGCCATCACCTGGACATCCACCAAATC 490 | |
| Db | 371 AGAAATTCCAGCTCCAGCTGGCCAGGCCATTCAGCTGAACCAATATGCTGAACCAACT 430 | 3134 AGCCCTCTCTAGAACCCCTGGAACTACTACAGCTACCCCCTCAACGCCATTG 3193 | |
| Qy | 3074 GCACAAACAGAGACCTGGACCTGAGAGATGCCATCACCTGGTCCCTCTGCCCCAGACTGCTACATCCACCAAATC 430 | 491 AGCCCTCTCTAGAACCCCTGGAACTACTACAGTACCCCTCAACGCCATTG 550 | |
| Db | 431 GCACAAACAGAGACCCCTGGAGACAGAGATGCCATCACCTGGACATCCACACTGAGATGCCATCACCTGGACATCCACCAAATC 490 | 3194 AGAACTGGGGAGCTGACTCTCTGCTACAGTGAGTGGAAAGGTCTCCAA 3250 | |
| Qy | 3134 AGCCCTCTCTAGAACCCCTGGAACTACTACAGCTACCCCCTCAACGCCATTG 3193 | 551 AGACTGGGGAGCTGACTCTCTGCTACAGTGAGTGGAAAGGTCTCCAA 607 | |
| Db | 491 AGCCCTCTCTAGAACCCCTGGAACTACTACAGTACCCCTCAACGCCATTG 550 | | |
| RESULT 14 | | | |
| AAL01390 | AAL01390 standard; cDNA; 572 BP. | | |
| XX | | | |
| AC | AAL01390: | | |
| XX | | | |
| DT | 21-NOV-2001 (first entry) | | |
| XX | | | |
| DB | Human reproductive system related antigen cDNA SEQ ID NO: 1391. | | |
| XX | | | |
| KW | Human; reproductive system related antigen; reproductive system disorder; cancer; gene therapy; ss. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |
| PN | WO20015320-A2. | | |
| XX | | | |
| PD | 02-AUG-2001. | | |
| XX | | | |
| PF | 17-JAN-2001; 2001WO-US01339. | | |
| XX | | | |
| PR | 31-JAN-2000; 2000US-017905. | | |
| PR | 04-FEB-2000; 2000US-0180628. | | |
| PR | 24-FEB-2000; 2000US-0184664. | | |
| PR | 02-MAR-2000; 2000US-0186350. | | |
| PR | 16-MAR-2000; 2000US-0189874. | | |
| PR | 17-MAR-2000; 2000US-0190076. | | |

PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-AUG-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225759.
 PR 14-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 05-SEP-2000; 2000US-0229513.
 PR 05-SEP-2000; 2000US-0230437.
 PR 05-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232124.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 03-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241888.
 PR 20-OCT-2000; 2000US-0241889.
 PR 20-OCT-2000; 2000US-0241890.
 PR 08-NOV-2000; 2000US-0241891.
 PR 08-NOV-2000; 2000US-0244677.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246530.
 PR 08-NOV-2000; 2000US-0246536.
 PR 08-NOV-2000; 2000US-0246537.
 PR 08-NOV-2000; 2000US-0246538.
 PR 08-NOV-2000; 2000US-0246539.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249220.
 PR 17-NOV-2000; 2000US-0249221.
 PR 17-NOV-2000; 2000US-0249224.
 PR 17-NOV-2000; 2000US-0249225.
 PR 17-NOV-2000; 2000US-0249226.
 PR 17-NOV-2000; 2000US-0249227.
 PR 17-NOV-2000; 2000US-0249228.
 PR 17-NOV-2000; 2000US-0249229.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251986.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251179.
 PR 06-DEC-2000; 2000US-0251856.
 PR 06-DEC-2000; 2000US-0251860.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251982.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678..
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-465570/50.
 P-PSDB; AAM95420.
 XX
 Isolated nucleic acid molecule encoding a reproductive system antigen -
 PT is used in preventing, treating or ameliorating a medical condition -
 XX
 PS Claim 1; SEQ ID NO 1391; 1297pp + Sequence Listing; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a coding sequence of the

CC invention.
 XX
 SQ Sequence 572 BP; 146 A; 159 C; 157 G; 106 T; 4 other;
 Best Local Similarity 9.4%; Score 411.2; DB 22; Length 572;
 Matches 492; Conservative 2; Mismatches 6; Indels 9; Gaps 7;
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0204677.
 PR 28-JUN-2000; 2000US-0214886.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217486.
 PR 14-JUL-2000; 2000US-0218290.
 PR 14-JUL-2000; 2000US-0215135.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225672.
 PR 14-AUG-2000; 2000US-0225248.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226229.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226888.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227099.
 PR 30-AUG-2000; 2000US-022894.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229513.
 PR 05-SEP-2000; 2000US-0230437.
 PR 05-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232030.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232337.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 27-SEP-2000; 2000US-0233064.
 PR 27-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234987.
 PR 25-SEP-2000; 2000US-0234988.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0236827.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.

RESULT 15

ABI96843 ABI96843 standard; cDNA; 572 BP.

AC ABL96843;

XX DT 21-JUN-2002 (first entry)

XX DE Human testicular antigen encoding cDNA SEQ ID NO: 511.

XX KW Human; testicular antigen; testes; cancer; metastasis; immune disorder; reproductive system disorder; urinary system disorder; gene therapy; cardiovascular disorder; respiratory disorder; neurological disorder; gastrointestinal disease; infection; cytostatic; gene; ss.

XX OS Homo sapiens.

XX PR 17-JAN-2001; 2001WO-US01329.

XX PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 02-AUG-2001.

XX WO200155317-A2.

PD

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Om nucleic - nucleic search, using sw model
Run on: January 6, 2004, 04:33:27 ; Search time 15537 Seconds
(without alignments)

11524.840 Million cell updates/sec
Title: US-10-054-691-1
Perfect score: 4377

Sequence: 1 atggggctggggccaggat.....tgcgcactgtggccctctag 4377
Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454013386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:
1: gb_ba: *
2: gb_htgb: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_p1: *
9: gb_pr: *
10: gb_r0: *
11: gb_sb: *
12: gb_sv: *
13: gb_un: *
14: gb_v1: *
15: em_ba: *
16: em_fan: *
17: em_hum: *
18: em_in: *
19: em_anu: *
20: em_on: *
21: em_or: *
22: em_over: *
23: em_patt: *
24: em_ph: *
25: em_p1: *
26: em_ro: *
27: em_sb: *
28: em_v1: *
29: em_vn: *
30: em_htgb_hum: *
31: em_htgb_inv: *
32: em_htgb_other: *
33: em_htgo_hum: *
34: em_htgo_mus: *
35: em_htgb_rod: *
36: em_htgb_mam: *
37: em_htgb_vrt: *
38: em_sy: *
39: em_htgo_other: *

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|-----------|----------------------------------|
| 1 | 4377 | 100.0 | 4377 | 6 | AX492941 | AX492941 Sequence |
| 2 | 3461.2 | 79.1 | 3648 | 6 | AX417817 | AX417817 Sequence |
| 3 | 3124 | 71.4 | 4612 | 4 | OCPHLIP | 212841 O. cuniculus |
| 4 | 2962 | 67.7 | 4608 | 10 | AF045454 | AF045454 Cavia por |
| 5 | 2687 | 61.4 | 4590 | 10 | D63648 | D63648 Rattus norvegicus |
| 6 | 2687 | 61.4 | 4613 | 6 | E13935 | E13935 Rat mRNA 5' |
| 7 | 1362 | 31.1 | 1835 | 6 | AX711962 | AX711962 Sequence |
| 8 | 1319.4 | 30.1 | 1790 | 9 | BC042674 | BC042674 Homo Sapiens |
| 9 | 970 | 22.2 | 2477 | 6 | AX713706 | AX713706 Sequence |
| 10 | 970 | 22.2 | 2477 | 9 | AK055428 | AK055428 Homo Sapiens |
| 11 | 576.2 | 13.2 | 608 | 6 | AX417821 | AX417821 Sequence |
| 12 | 372.2 | 8.5 | 583 | 6 | AX417820 | AX417820 Sequence |
| 13 | 322.2 | 7.4 | 820 | 10 | BC03306 | BC03306 Mus musculus |
| 14 | 251 | 5.7 | 265 | 6 | AX417822 | AX417822 Sequence |
| 15 | 237.4 | 5.4 | 446 | 6 | AX417823 | AX417823 Sequence |
| 16 | 205 | 4.7 | 43543 | 6 | AX711964 | AX711964 Sequence |
| 17 | 205 | 4.7 | 180465 | 9 | AC074011 | AC074011 Homo Sapiens |
| 18 | 177.2 | 4.0 | 16595 | 2 | AC022076 | AC022076 Homo Sapiens |
| 19 | 120.6 | 2.8 | 1672 | 3 | AY069174 | AY069174 Drosophila melanogaster |
| 20 | 109 | 2.5 | 10926 | 9 | AC093164 | AC093164 Homo Sapiens |
| 21 | 107.8 | 2.5 | 357 | 6 | AK314612 | AK314612 Sequence |
| 22 | 107.2 | 2.4 | 50250 | 2 | AC017341 | AC017341 Drosophila melanogaster |
| 23 | 107.2 | 2.4 | 17283 | 3 | AC009383 | AC009383 Drosophila melanogaster |
| 24 | 107.2 | 2.4 | 299903 | 3 | AE003514 | AE003514 Drosophila melanogaster |
| 25 | 94.6 | 2.2 | 101340 | 2 | AC018104 | AC018104 Drosophila melanogaster |
| 26 | 94.6 | 2.2 | 190668 | 3 | AC007419 | AC007419 Drosophila melanogaster |
| 27 | 94.6 | 2.2 | 260731 | 3 | AE003611 | AE003611 Drosophila melanogaster |
| 28 | 85.6 | 2.0 | 21253 | 2 | AC012372 | AC012372 Mus musculus |
| 29 | 82.8 | 1.9 | 214038 | 2 | AC141490 | AC141490 Rattus norvegicus |
| 30 | 81.6 | 1.9 | 214038 | 2 | AC141490 | AC141490 Rattus norvegicus |
| 31 | 75.6 | 1.7 | 21253 | 2 | AC102372 | AC102372 Mus musculus |
| 32 | 70.4 | 1.6 | 13246 | 9 | AC125617 | AC125617 Homo Sapiens |
| 33 | 68.6 | 1.6 | 684 | 4 | CFA234375 | AJ234375 Canis familiaris |
| 34 | 68.2 | 1.6 | 124020 | 9 | AF429315 | AF429315 Homo sapiens |
| 35 | 63.2 | 1.4 | 124020 | 9 | AF429315 | AF429315 Homo sapiens |
| 36 | 61.6 | 1.4 | 631 | 11 | BV055548 | BV055548 S21P6131 |
| 37 | 59.8 | 1.4 | 44237 | 3 | AC024847 | AC024847 Caenorhabditis elegans |
| 38 | 59.8 | 1.4 | 267118 | 2 | AC006889 | AC006889 Caenorhabditis elegans |
| 39 | 58 | 1.3 | 198736 | 5 | AL929023 | AL929023 Zebrafish |
| 40 | 52.6 | 1.2 | 20706 | 6 | AX655393 | AX655393 Sequence |
| 41 | 52.6 | 1.2 | 123789 | 2 | AC110674 | AC110674 Tetrahymena thermophila |
| 42 | 52.2 | 1.2 | 2000 | 6 | AX655393 | AX655393 Sequence |
| 43 | 49.2 | 1.1 | 232304 | 2 | AC111030 | AC111030 Mus musculus |
| 44 | 47.8 | 1.1 | 34305 | 3 | CW002B12 | CW002B12 Caenorhabditis elegans |
| 45 | 47.2 | 1.1 | 212127 | 5 | BX248515 | BX248515 Zebrafish |

ALIGNMENTS

RESULT 1
AX492941
LOCUS Sequence 1 from Patent WO20059328.

DEFINITION ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo;

Turner Jr.,C.A., Miranda,M. and Yu,X.

Human lissase and polynucleotides encoding the same

Patent: WO 20059328-A 01-AUG-2002;

| | | | |
|----|------|---|------|
| Db | 3061 | GAACCACTGGAGCAAACAGAGACCTGGACTGAGGCAGAGATGCCCATCACGT | 3120 |
| Qy | 3121 | CCCACTCGAATGAGGCCCTCCTGAGAACCCCTCGAAATACTAACCTPACCCATC | 3180 |
| Db | 3121 | CCCACTCGAATGAGGCCCTCCTGAGAACCCCTCGAAATACTAACCTPACCCATC | 3180 |
| Qy | 3181 | AAGCCACGCAATTGAGAACTGGGACTGTGACTTCCTCTGTGAGACTGAGGCTTCAT | 3240 |
| Db | 3181 | AAGCCACGCAATTGAGAACTGGGACTGTGACTTCCTCTGTGAGACTGAGGCTTCAT | 3240 |
| Qy | 3241 | AGTGTTCACCTCTGTCACCGCTGCCCTGAGACATCCAGTT | 3300 |
| Db | 3241 | AGTGTTCACCTCTGTCACCGCTGCCCTGAGACATCCAGTT | 3300 |
| Qy | 3301 | GGTGACTCTCGACTACAGCTGAGCTGAGCTGACCTACCTCCACA | 3360 |
| Db | 3301 | GTTGACTCTCGACTACAGCTGAGCTGACCTACCTCCACA | 3360 |
| Qy | 3360 | TCTTGAGGGACTCTTGAGCATGGGGATGGAACTGGACTCACCCA | 3420 |
| Db | 3360 | TCTTGAGGGACTCTTGAGCATGGGGATGGAACTGGACTCACCCA | 3420 |
| Qy | 3421 | CIGCCCAACATCTGAGAAAGTCAACCCTTACCTCTCTGTCTCTACAGCACCTGG | 3480 |
| Db | 3421 | CIGCCCAACATCTGAGAAAGTCAACCCTTACCTCTCTGTCTCTACAGCACCTGG | 3480 |
| Qy | 3481 | GAGGAGACAGGAGACTTAATGTGGAGTTCACGATCATTCACTAACGCTGTCCC | 3480 |
| Db | 3481 | GAGGAGACAGGAGACTTAATGTGGAGTTCACGATCATTCACTAACGCTGTCCC | 3480 |
| Qy | 3491 | CAGCCCTGGACCTGGTAGAGGAATGAAACAGCCGACATCACTGGAGAAAGAC | 3600 |
| Db | 3491 | CAGCCCTGGACCTGGTAGAGGAATGAAACAGCCGACATCACTGGAGAAAGAC | 3600 |
| Qy | 3541 | CAGGCCACTTGCCAGGAATATGTCACATCCACAGGGCTGGACATCTGAGAAGAT | 3520 |
| Db | 3541 | CAGGCCACTTGCCAGGAATATGTCACATCCACAGGGCTGGACATCTGAGAAGAT | 3520 |
| Qy | 3560 | TGAAAGTTGGCACACTCTTCAATGGGTCAACGACTTGTCTCATTAATGTGAGAATCG | 3660 |
| Db | 3560 | TGAAAGTTGGCACACTCTTCAATGGGTCAACGACTTGTCTCATTAATGTGAGAATCG | 3660 |
| Qy | 3661 | GAGGCCACTTGCCAGGAATATGTCACATCCACAGGGCTGGACATCTGAGAAGAC | 3660 |
| Db | 3661 | GAGGCCACTTGCCAGGAATATGTCACATCCACAGGGCTGGACATCTGAGAAGAC | 3660 |
| Qy | 3721 | GAGGACTCCAGGGCTTGTCAAGTGTGGAGGTATGGACTGGCTAGCCGTAC | 3780 |
| Db | 3721 | GAGGACTCCAGGGCTTGTCAAGTGTGGAGGTATGGACTGGCTAGCCGTAC | 3780 |
| Qy | 3781 | CAGGGCAAGGGGAATGTGCCATGTCGACAGCTGAGAACACTGCACTGCTCAGA | 3840 |
| Db | 3781 | CAGGGCAAGGGGAATGTGCCATGTCGACAGCTGAGAACACTGCACTGCTCAGA | 3840 |
| Qy | 3841 | CACTCGCAAGTCCCTGGAGAGCAAGACTGAAAGTGAAGCTGGAACTTCCAGAT | 3900 |
| Db | 3841 | CACTCGCAAGTCCCTGGAGAGCAAGACTGAAAGTGAAGCTGGAACTTCCAGAT | 3900 |
| Qy | 3901 | GGCATCTCAGTTCTCCACTGGCACCAATACACAGCGCTGAGACTTCAGA | 3960 |
| Db | 3901 | GGCATCTCAGTTCTCCACTGGCACCAATACACAGCGCTGAGACTTCAGA | 3960 |
| Qy | 3961 | GTGCGACCTTCTCCAAACACACTCACCCACAGCAAGAGAGGGACACTGACCTC | 4020 |
| Db | 3961 | GTGCGACCTTCTCCAAACACACTCACCCACAGCAAGAGAGGGACACTGACCTC | 4020 |
| Qy | 4021 | ACCTCTCTCGAGGACTGTGTTACTCTCAGACCGCGGAGATGGCTTG | 4080 |
| Db | 4021 | ACCTCTCTCGAGGACTGTGTTACTCTCAGACCGCGGAGATGGCTTG | 4080 |
| Qy | 4081 | GAACCTGGAAACACTGAGGACCTGGCCAGACTTCCACACTTCAC | 4140 |
| Db | 4081 | GAACCTGGAAACACTGAGGACCTGGCCAGACTTCCACACTTCAC | 4140 |
| Qy | 4141 | CACAGCGAGCAACTCAGTGCCCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGG | 4200 |
| Db | 4141 | CACAGCGAGCAACTCAGTGCCCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGG | 4200 |

QY 4201 AACAGCCGATTGCTCCAGACCGCTGAGAAGAGCCCCGAGCTACTGGCTTC 4260
Db 4201 AACAGCCGATTGCTCCAGACCGCTGAGAAGAGCCCCGAGCTACTGGCTTC 4260
QY 4261 CGATGGAGCAGCAGGAGTGGCCCTGTGGTGGCATCATGGACAGTGGTCTGGTC 4320
Db 4261 CGATGGAGCAGCAGGAGTGGCCCTGTGGTGGCATCATGGACAGTGGTCTGGTC 4320
QY 4321 AGGAGAGTGGCGGAGGAGATCCCATAGGCCCTCGACTGTGGCCCTAG 4377
Db 4321 AGGAGAGTGGCGGAGGAGATCCCATAGGCCCTCGACTGTGGCCCTAG 4377

RESULT 2

AX417817 AX417817 Sequence 1 from Patent WO20021161. DNA linear PAT 18-JUN-2002
LOCUS ACCESSION VERSION KEYWORDS SOURCE ORGANISM
JOURNAL
FEATURES source
BAYER AG (DE) location/Qualifiers 1. .3448 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="Lxon:9606"
BASE COUNT 899 a 1012 c 947 g 770 t
ORIGIN

Query Match Best Local Similarity 97.5%; Pred. No. 0; Mismatches 33; Indels 57; Gaps 2;
Matches 3558; Conservative

QY 775 GAAGCTCTGAAACAGCCTCTGGCTCCAGCAAGTACAGTGAGCAGGCTCACCCTG 834
Db 1 GAAGCTCTGAAACAGCCTCTGGCTCCAGCAAGTACAGTGAGCAGGCTCACCCTG 60
QY 835 GTTTCCAGCCTCTCTATGAGACCAACCATCTCTACACTGGAGGACCCGACTC 894
Db 61 GTTTCCAGCCTCTCTATGAGACCAACCATCTCTACACTGGAGGACCCGACTC 120
QY 895 CAGGATTCACAGCTGCCTGCACTCTGGATAGTAGATGAGCTGGAGCAACAGGAG 954
Db 121 CAGGATTCACAGCTGCCTGCACTCTGGATAGTAGATGAGCTGGAGCAACAGGAG 180
QY 955 AAAGATGAGCCATTGAGCTTAACACGGAGCCATTGAGTCCTCTGGAGG 1014
Db 181 AAAGATGAGCCATTGAGCTTAACACGGAGCCATTGAGTCCTCTGGAGG 240
QY 1015 CCTTATCTTCTGCTACAGAACAGCAACTCTGAGCAGTGGAGCTGAGAACCCAAAGC 1074
Db 241 CCTTATCTTCTGCTACAGAACAGCAACTCTGAGCAGTGGAGCTGAGAACCCAAAGC 300
QY 1075 AAGCTTG-----AGGTAAGAGA 1092
Db 301 AAGCTTGAGAAAATCCATTCTTATTCCGTGTCATTCAAGTAAGAGA 360
QY 1093 GGAGGGAAATCAGATGCTGCTGACAAGACCCCTCGATACTGGTCCACCTCAGTC 1152
Db 361 GGAGGGAAATCAGATGCTGCTGACAAGACCCCTCGATACTGGTCCACCTCAGTC 420
QY 1153 AGGTGAAGCCGCTGACATCAAGCTAATTGGAGCCCTGGTGAATGAAATTGCTCC 1212
Db 421 AGGTGAAGCCGCTGACATCAAGCTAATTGGAGCCCTGGTGAATGAAATTGCTCC 480

QY 1213 AATGGGGGGGGGGTCAACACCTGGAAACCTCTGGCTGACTCAGTACCGAGGCTG 1272
Db 481 AATGGGGGGGGTCAACACCTGGAAACCTCTGGCTGACTCAGTACCGAGGCTG 540
QY 1273 TCTGGAGCGCTGGCGAGATGAGAACATCGCACCGTACACCCCTGGCAGACATCCC 1332
Db 541 TCTGGAGCGCTGGCGAGATGAGAACATCGCACCGTACACCCCTGGCAGACATCCC 600
QY 1333 CGGAATTCAACCTCTCTGAGGGCTCTCTGTGACTGGAAAGANACAGCTCT 1392
Db 601 CGGAATTCAACCTCTCTGAGGGCTCTCTGTGACTGGAAAGANACAGCTCT 660

QY 1393 AATGCCCTCTTAAACAGCTGGCGAGGAGCCGAGCTGAGGATCTACCTGCGCC 1452
Db 661 AATGCCCTCTTAAACAGCTGGCGAGGAGCCGAGCTGAGGATCTACCTGCGCC 720

QY 1453 ACGAGGCTGGTGGACCTGAGAAGTACACTTCAGGAAGACTGAG 1512
Db 721 AGGAGGCTGGTGGACCTGAGAAGTACACTTCAGGAAGACTGAG 780

QY 1513 ATATACCTGTATAGGCCCATACCTCTGTGATTCGCAATGATCTGGCC 1572
Db 781 ATATACCTGTATAGGCCCATACCTCTGTGATTCGCAATGATCTGGCC 840

QY 1573 TATTCCTCCAGACTTCAGAGAACATGAGAAGTACACTTCAGGAAGACTGAG 1632
Db 841 TATTCCTCCAGACTTCAGAGAACATGAGAAGTACACTTCAGGAAGACTGAG 900

QY 1633 GTTCTCGGGCATTTGAACTCTGTCAGGTCAGGTTGCTGAGATGTCACCTGGAGCTG 1692
Db 901 GTTCTCGGGCATTTGAACTCTGTCAGGTCAGGTTGCTGAGATGTCACCTGGAGCTG 960

QY 1693 TACAGGGAGAAAGTACTCTGCCCCAGGAATGATCTCTAGTGTCTGTCAGGTCACCTGGAGCTG 1752
Db 961 TACAGGGAGAAAGTACTCTGCCCCAGGAATGATCTCTAGTGTCTGTCAGGTCACCTGGAGCTG 1020

QY 1753 CTAGAAGTTGATGATACTCAACAGAACTTGCTACCTCATGAAATCAGAGTT 1812
Db 1021 CTGAGTTGATGATACTCAACAGAACTTGCTACCTCATGAAATCAGAGTT 1080

QY 1813 CAGGAGAGACCCACCAACTGATGAGCTGGCGATATGACACAAAGGAAGTT 1872
Db 1081 CAGGAGAGACCCACCAACTGATGAGCTGGCGATATGACACAAAGGAAGTT 1140

QY 1873 GTGTTGGCAAGCCGTTTGAACCTGGACATGCCAACGACTGGAGGATTCCT 1932
Db 1141 GTGGTGGCAAGCCGTTTGAACCTGGACATGCCAACGACTGGAGGATTCCT 1200

QY 1933 GACAACTTCTCGCTCTGACTGTTCACTTCAGGAGTCACCTCCGAGCA 1992
Db 1201 GACAACTTCTCGCTCTGACTGTTCACTTCAGGAGTCACCTCCGAGCA 1260

QY 1993 GCGAGTGTCTCTGGAAACATTGCTGGAGCTGTTGGCCAGAGACACTGCTATA 2052
Db 1261 GCGAGTGTCTCTGGAAACATTGCTGGAGCTGTTGGCCAGAGACACTGCTATA 1320

QY 2053 TTGAAACAAAGTCAATCATGCTGGAGCCAGCTTCCAGCGTTCCTGAGGACT 2112
Db 1331 TTGAAACAAAGTCAATCATGCTGGAGCCAGCTTCCAGCGTTCCTGAGGACT 1380

QY 2113 AAGACAG-----CATGGAGGTCATGGACCTGGCTGGCCATGAGGGAC 2157
Db 1381 AGATTCGGATTGTTCTTCTGGGTCATGGACCTGGCTGGCCATGAGGGAC 1440

QY 2158 AGGCCCCCTTCGCTCTGACCTCTGGAGCTGCTGAGACCTGCTGAGACATCCAA 2217
Db 1441 AGGCCCCCTTCGCTCTGACCTCTGGAGCTGCTGAGACCTGCTGAGACATCCAA 1500

QY 2218 GTGTTGGCTGCTGGAGATCTCTGACCTCTGGAGCTGCTGAGACCTGCTGAGAC 2277
Db 1501 GTGTTGGCTGCTGGAGATCTCTGACCTCTGGAGCTGCTGAGACCTGCTGAGAC 1560

| | | | | | | | |
|----|------|---|------|----|------|---|------|
| QY | 2278 | GACCTCCCGGAGTCACCAACAGTACGGGACTGTATAAGTGAGGAGGGAGGGC | 2337 | Db | 2641 | TCGAGGGACTCTTGGAGCATGGGGATGGGAACTTGAGACTCACCC | 2700 |
| Db | 1561 | GACCTCCCGGAGTCACCAACAGTACGGGACTGTATAAGTGAGGAGGGAGGGC | 1620 | QY | 3418 | AACATGCCAACATTCTGAAGAGTCAACCTTACCTCTGGCTTCCTACAGACC | 3477 |
| QY | 2338 | TCCCTGGAGAAGTGACCAACCTTAACATATCTTCGGGTTAACAGAACCTACA | 2397 | Db | 2701 | AACATGCCAACATTCTGAAGAGTCAACCTTACCTCTGGCTTCCTACAGACC | 2760 |
| Db | 1621 | TCCCTGGAGAAGTGACCAACCTTAACATATCTTCGGGTTAACAGAACCTACA | 1680 | QY | 3478 | TCGGAGGGACAGCAGGACTAAATGTCGAGGGAAAGGGCAGAGTCACTCCA | 3537 |
| QY | 2398 | GCTCTACGGCGTGGCAACGGTGAATGACAGAATGCTCAATCAGCTGTT | 2457 | Db | 2761 | TCGGAGGGACAGCAGGACTAAATGTCGAGGGAAAGGGCAGAGTCACTCCA | 2820 |
| Db | 1681 | GCTCTACGGCGTGGCAACGGTGAATGACAGAATGCTCAATCAGCTGTT | 1740 | QY | 3538 | GCCTAGGCTGGACCTTAACCTTCTGGGTTAACAGGATCTGCTTCTGGAGAA | 3597 |
| QY | 2458 | CCGGGACAAAGCTGAGGATTTATGAGCCAGTCAAGTCACTGCTGAGATGAA | 2517 | Db | 2821 | GCCGAGGCTGGAGCTGAGGATGAGGAGATGAGAAACAGCCGACATCCTGGAGAA | 2880 |
| Db | 1741 | CCGGGACAAAGCTGAGGATTTATGAGCCAGTCACTGCTGAGATGAA | 1800 | QY | 3598 | GACTGGAGCTGGTCAACTCTCATGGGGTCAACACTGTGCTTACTGTGAGAT | 3657 |
| QY | 2518 | GATGATCATAGATAAATTCTCATGAGACTGAGAAGGTCATACAGTGTGAGG | 2577 | Db | 2881 | GACTGGAGCTGGTCAACTCTCATGGGGTCAACACTGTGCTTACTGTGAGAT | 2940 |
| Db | 1801 | GATGATCATAGATAAATTCTCATGAGACTGAGAAGGTCATACAGTGTGAGG | 1860 | QY | 3658 | CCGGAGGCCACTTGSCCAGGAATATGTCGACATCCACAGGCCCCTGGACATCCTC | 3717 |
| QY | 2578 | ASGGATTATGACTACTGCAAGATTCGAACTGTGATTCAGCTGAGATGAA | 2637 | Db | 2941 | CCGGAGGCCACTTGSCCAGGAATATGTCGACATCCACAGGCCCCTGGACATCCTC | 3000 |
| Db | 1861 | ASGGATTATGACTACTGCAAGATTCGAACTGTGAGATGAACTTGTGAC | 1920 | QY | 3718 | TCTGAGGAGCTCCAAGGGCTTGTGCAAGGAGCATGAGCTGCTGAGCTGCTG | 3777 |
| QY | 2638 | CATCTCCCAAATCTTGGACTGCTCTGACATAGAGGTCGAGGAGTCCTGTCACCTC | 2697 | Db | 3001 | TCTGAGGAGCTCCAAGGGCTTGTGCAAGGAGCATGAGCTGCTGAGCTGCTG | 3060 |
| Db | 1921 | CATCTCCCAAATCTTGGACTGCTCTGACATAGAGGTCGAGGAGTCCTGTCACCTC | 1980 | QY | 3778 | TACAGGGCCAAAGGGGAAATGTGCACTGGGCACTCAGGAGCTGCTGAGCTGCTC | 3837 |
| QY | 2698 | GTCGACTCTGAAACCCACTATCATGGGCAAGGGTCTCTGGAAACCCAGACAATG | 2757 | Db | 3061 | TACAGGGCCAAAGGGGAAATGTGCACTGGGCACTGAGCTGAGCTGCTGAGCTGCTC | 3120 |
| Db | 1981 | GTCGACTCTGAAACCCACTATCATGGGCAAGGGTCTCTGGAAACCCAGACAATG | 2040 | QY | 3838 | AGACATCGAAAGCTCCCTGGAGAAAGCAAGACTGAGAAGTGAACCTGGACCTCCG | 3897 |
| QY | 2758 | CCAGTGCAGCAACCCAGCGTTAGTGTACTCTGACCTGCGCGAGACTCCAA | 2817 | Db | 3121 | AGACATCGAAAGGGGAAATGTGCACTGGGAGAGACTGAGAAGTGAACCTGGAC | 3180 |
| Db | 2041 | CCAGTGCAGCAACCCAGCGTTAGTGTACTCTGACCTGCGGGAGACTCCAA | 2100 | QY | 3898 | CATGGCATCTCAGTTCTTCACTGSGCCAAATACACAGCGAGGAGCTGGGTT | 3957 |
| QY | 2818 | GGCTGAGCCGCTGGAGGGCTTCAGGGAGGCTACGGAGCATCGGGAGCTG | 2877 | Db | 3181 | CATGGCATCTCAGTTCTTCACTGSGCCAAATACACAGCGAGCTGGGTT | 3240 |
| Db | 2101 | GAAGCTGAGCCGCTGGAGGGCTTCAGGGAGGCTACGGAGCATCGGGAGCTG | 2160 | QY | 3959 | GTGGTGGAGCCCTTCCTCCAAACACACTCAGGGAGCTGAGAGGGGAACTGAC | 4017 |
| QY | 2878 | GSGTCAGGCGGTATGAGACGGAGGAGCTCTCTGGGTTGTGAGCCCTCTCCAG | 2937 | Db | 3241 | GTGGTGGAGCCCTTCCTCCAAACACACTCAGGGAGGAGCTGAGAGGGGAACTGAC | 3300 |
| Db | 2161 | GSGTCAGGCGGTATGAGACGGAGGAGCTCTGGGTTGTGAGCCCTCTCCAG | 2220 | QY | 4018 | CTCACCTTCTCCAGGAGACTGTTTCACCTCTCAGAGCCGGGGATGCCAGATGG | 4077 |
| QY | 2938 | ACATCCAGCTCCTGRCCTGCGGGATGGCTCCAGATACCTCTTGTGAGAC | 2997 | Db | 3301 | CTCACCTTCTCCAGGAGACTGTTTCACCTCTCAGAGCCGGGGATGCCAGATGG | 3360 |
| Db | 2221 | ACATCCAGCTCCTGRCCTGCGGGATGGCTCCAGATACCTCTTGTGAGAC | 2280 | QY | 4078 | ATGGCACTCTGAGAACACATGCTGGACACAGCTGGGCGCAAGACTACCTCCAAACACTC | 4137 |
| QY | 2998 | TGATCCACCCAAATCAGAACATTCACTCCAGCTGGCAGGCCCTTGGACCATATG | 3057 | Db | 3361 | ATGGCACTCTGAGAACACATGCTGGACACAGCTGGGCGCAAGACTACCTCCAAACACTC | 3420 |
| Db | 2281 | TGATCCACCCAAATCAGAACATTCACTCCAGCTGGCAGGCCCTTGGACCATATG | 2340 | QY | 4138 | ACCCACGCCGAGCCAAACTCAAGTGGCCCTTCCTGGAGAGCCCTTAACACCTG | 4197 |
| QY | 3058 | CTGGAACCACTTGGAGAACAGAGAACCTGGACCTGGAGAGCATGCCATCACC | 3117 | Db | 3421 | ACCCACGCCGAGCCAAACTCAAGTGGCCCTTCCTGGAGAGCCCTTAACACCTG | 3480 |
| Db | 2341 | CTGGAACCACTTGGAGAACAGAGAACCTGGACCTGGAGAGCATGCCATCACC | 2400 | QY | 4198 | CGAACACGGCGATGCTCCAGACAGGCTGAAGAGGCCCGAGGTCTACTGGCT | 4257 |
| QY | 3118 | TGTGCTCCACTCAAGATGAGAACCCCTGGAGAATGACTCTGCTGAGCTG | 3177 | Db | 3481 | CGAACACGGCGATGCTCCAGACAGGCTGAAGAGGCCCGAGGTCTACTGGCT | 3540 |
| Db | 2401 | TGTGCTCCACTCAAGATGAGAACCCCTGGAGAATGACTCTGCTGAGCTG | 2460 | QY | 4258 | CTCCCACTGGAGGGAGGGCTGCTGAGGAGATCTCTCAATGAGCTGGGAGCT | 4317 |
| QY | 3238 | ATAGTGTCCAACTCTGTCACCCAGCTCCACCAACAGAGCATCAAAGTGGGCC | 3297 | Db | 3541 | CTCCCACTGGAGGGCTGCTGAGGAGATCTCTCAATGAGCTGGGAGCT | 3600 |
| Db | 2521 | ATAGTGTCCAACTCTGTCACCCAGCTCCACCAACAGAGCATCAAAGTGGGCC | 2580 | QY | 4318 | TCCAGGAGGTGGAGGGAGATCTCTCAATGAGCTGGGAGCT | 4365 |
| QY | 3298 | CTGGGTGACTCTGTGACTACACAGCTGGAGCTGACCAACACTCAGTGTGAC | 3357 | Db | 3601 | TCCAGGAGGTGGAGGGAGATCTCTCAATGAGCTGGGAGCT | 3648 |
| Db | 2581 | CTGGGTGACTCTGTGACTGTGAGTGTGGAGGAGTCGACCAACACTCAGTGTGAC | 2640 | QY | 3358 | ACATCTTGAGGGGACTCTCTGGAGGATTGGAGGGATGGGAACTTGGAGACTCACACC | 3417 |

RESULT 3

OCPHILP

LCUS

DEFINITION

O_cuniculus mRNA for phospholipase.

ACCESSION

Z12841

VERSION Z12841.1 GI:1689

SOURCE
Oryctolagus cuniculus (rabbit)

ORGANISM
Oryctolagus cuniculus

REFERENCE
Bukaryota; Metzca; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

JOURNAL
J. Biol. Chem. 268 (17), 12901-12911 (1993)
MEDLINE
93286138

PUBLISHED
8509/24
2 (bases 1 to 4612)

AUTHORS
Boll,W.

TITLE
Submitted (25-JUN-1992) WERNER BOLL, Dept. of Biochemistry 11,
Swiss Federal Institute of Technology, ETH Zentrum,
Universitätstrasse 16, Zuerich, CH-8092, Switzerland

FEATURES
Source
Direct Submission
Location/Qualifiers
1. .4612 "Oryctolagus cuniculus"
/organism="Oryctolagus cuniculus"
/mol_type="mRNA"
/strain="New Zealand White"
/ab_xref="taxon:9986"
/clone="padrab-113"
/cell_type="Epithelial cell"
/tissue_type="Small intestine"
/dev_stagex="Adult"
25. .4401
codon_start=1
/product="Phospholipase"
/protein_id="CAA78303.1"
/db_xref="GI: 1690"
/translation="MAIKPSVFLGLQPLPGRGADQIQTSSGGNTLQLQWPESLKTP
PPPCDPKTKLAEVPSSEVHSLSPLPSDKRVAATGAACTTCAGATTAACAGCCATTGCG
ACMGVYVUTLSDITIGRSPLAMPLPSDKRVAATGAACTTCAGATTAACAGCCATTGCG
EHDWKLINVFVNTSQFCFCFSAQQKGAVLGGNDKLIITDYLQOQVVKATVNLUDS
ELAAFSKWRQGQLSPAEPCRLTRETSQLTQVLYLWDLSLASSKNTQSP
AVUFOPEPFYESSLUSALAEPPPLCPLTRETSQLTQVLYLWDLSLASSKNTQSP
TQESPFLTYRNQGLTRVSQIQGKLEKREGTEIREPKDFKDSVTSVHLKPADK
VIGAMGUDSLTAGNSQSGPNLIDVQYRGSWSVNGDQNTSTVTHLNTREIFPS
LOGFSVUGTGRETSQARNOAVAGARDYLQIPOAQLVALMDNTRINFODWKIV
FIGGNDLICDFCNDPVRISQNTDNCIALDILHAEPPRAFLNVLKULISKRELIQ
ETRKVSCPRMILRSLCPVLTQFDNSTASLJETKEYQVTOQDILSGRDVTQDPT
VUQOPFEPEKVNMKPTQGCLPDNSFAEFSKOHAAHAAASALWNRMPEPTKREKEP
NDFEGAVINTCENQVWQFLSTKNSVOGGFTWLPCROSSPSASPPSVAHLKPADQ
VALAQLSITAGIGIGKSENDSDGTYRQYRGYSVSGEDDSGSDNVTUPTLNUFNPNTL
IGASDLCIQCUTSNSLYKAANFTHLRKDALAHREVERALVNLVDPMPSVTKQVLG
NFDKCPVQOQASALCNCVSPRNSYELARLAQAOVOSSERELVSGSYRTDESV
VLOQPFPSIOLPVQDGRSLPDRAPCVPHQKQFHSQSLRALWMPLGGKD
DITAATTCPTCNEPPLTRFRNSDYPYPSRPAVENGMSDFLCTAMMASRGVNPSTE
LQPGDIIKWAQALGDSLUTAWGARPSNSDPPWFGSWSVSGGDGALETHTLPLNK
IGASDLCIQCUTSNSLYKAANFTHLRKDALAHREVERALVNLVDPMPSVTKQVLG
NFDKCPVQOQASALCNCVSPRNSYELARLAQAOVOSSERELVSGSYRTDESV
VLOQPFPSIOLPVQDGRSLPDRAPCVPHQKQFHSQSLRALWMPLGGKD
DITAATTCPTCNEPPLTRFRNSDYPYPSRPAVENGMSDFLCTAMMASRGVNPSTE
VAGRGMMCRPRDPPPLSITGL"

sig_peptide
polyA_signal
polyA_site
BASE COUNT ORIGIN

Query Match 71.4%; Score 3124; DB 4; Length 4612;
Best Local Similarity 82.6%; Pred. No. 0; Mismatches 750; Indels 9; Gaps 3;
Matches 3613; Conservative 0; MisMatches 750; Indels 9; Gaps 3;

1 ATGGGGCTGGGCCGGCAAGGCAATTCTCTCTGGAGCTGCTCTGCTGGGCCAGGGCC 60

Db 25 ATGGCTCTGAGCAGGGCTCTCTGGAGCTGCTCCCTGGGAGGTC 84

Db 61 CCTAGATCCATRACTCTCTGAAAGGTACTTGGAGGSCAGCTATGSCAGAGGC 120

Db 85 GATCAGATCCAACTCTCTGAGAACAGACATGGAGCTGCTCTGGAGGCCAGC 144

Db 121 CTGAGAACTTCATCCATGAAACCATAATTAGGAGTAAATGCTCAA 180

Db 145 CTGAGATTTTCATCCATGAAAGACATGGAGGCCAGC 240

Db 181 TCTGTTCACTCTGAGCTCTGATTAATTTGGCGCCATGGCAATCGAA 240

Db 205 TCACTCACTCTGAGGCTTCAGCATTAATTGTCAGACCCATGGCAATGCAA 264

Db 241 ATTCTCCAGACCCAGGGAGGCGATCTGGAGGAGAGATGGATG--AAAGCCA 297

Db 265 ACTGCTCCAGATCAGGGCGAGCAGCTAGAGGAGCAGGGAGGAGGCCA 324

Db 298 CAGGAGGTGAGGGAGGAGTACACCTCTTCAGACATACAGATATTGCT 357

Db 325 GAGGAGGTGAGGGAGGAGTACACCTCTTCAGACATACAGATATTGCT 384

Db 358 TCTGTTCAATGCTCTGAGCTGAGACTGAACTGAGAGAGTCATCCACAGATATCGCCGTTCACT 384

Db 385 TCTGCTCTGATGCGCTGTC--CTGAGACAGACTGTTACCTCGAGACTGCA 417

Db 418 GACTGTGAGTCACTGAGCTGAGCTGAGAGAGTCATCCACAGATATCGCCGTTCACT 417

Db 442 GACTGTGAGTCACTGAGCTGAGCTGAGAGAGTCATCCACAGATATCGCCGTTCACT 501

Db 478 TTCGATTTGACTGAGACTCAATGCTGTTCTCACTAACTGAGCTGTTACAGCCAGTTACCTG 537

Db 502 TTGACACATGACTGAGCTCATCACCTGTTCTCAACACAGCCAGTTACCTG 561

Db 529 TGCGCTCTGCTCAACAGAACTGGCTCTGCTGGGGCATGACAAGGACCTTCCC 597

Db 562 TGCGCTCTGCTCAACAGAACTGGCTCTGCTGGGGCATGACAAGGACCTTCCC 621

Db 598 CTGGACTACTCTGAGGGAGGCTCCAGAGCTTGTAACTCTGGAGACTCTCTGAG 657

Db 622 CTGGACTACTCTGAGGGAGGCTCCAAAGGGTGCAACTCTGGAGATCTCTGAG 681

Db 658 GTTCGAGGGCTCTGAGATCACCGCATCTGGCTCAGGCCAGGCCAGCCCTGT 717

Db 682 TTGGCAGCATCTCTCGCTGGCTCAGGGGTCACTGCAAGCCAGGCCAGCCCTGT 741

Db 718 ATGGCTCTGAGGAGACCCCGCTGCCAACGGGGTCACTGCAAGCCAGGCCAGCCCTGT 777

Db 742 AGGTGCTTAAAGGAAACCTCCAGCTGACAAAGTCTCTGACAGCTGGCTTACCTGAA 801

Db 778 GCTCTGGAAACGGCTCTGAGGAGCTGCCAACGGGGTCACTGCAAGCCAGGCCAGCCCTGT 837

Db 802 GCTCTGGAAACGGCTCTGAGGAGCTGCCAACGGGGTCACTGCAAGCCAGGCCAGCCCTGT 861

Db 838 TCCAGGCCTTCTCTGAGAACCCATCTTACACTCTGAGGAGCCCACTCG 897

Db 862 TTCCAGGCCTTCTCTGAGAACCCATCTGAGGAGCTGCCAACGGGGCCCACTCCG 921

Db 898 GATCTTACACCTGGCTGGCATCTCTGAGTAGGATGATGAGCTGAGCCAGCAGGAA 957

Db 922 GACCCACACACTGGCCCTGACTCTGAGAACAGATGATGAGCTGCCAACGGCC 981

Db 958 GATCAGCCATTGAGTGTAAACAGGGAGGCAATGAGTGTCTGAGGCCAGGCC 1017

Db 1042 TACCTTACCTCTGAGAACAGGGCCAGCTGAGCTGCCAACGGGG 1101

Db 1078 CTGAGGTTAGAGGAGGAGGAGGAATGAGTGTGAGCTGCCAACGGAGGCC 1041

Qy 1018 TATCTGTTCACTGAGTGTAAACAGGGAGGCAATGAGTGTCTGAGGCCAGGCC 1077

| | | | | | | | |
|----|------|--|------|----|------|---|------|
| Db | 1102 | CTCGAGGGAGAGAGGAGGAGGAATCGCT3TCCGACAAGAACCCCTCCACTCGGT | 1161 | Oy | 2218 | GTTGTCGCTGCTGGGGATTCTCTGACGCCCTGGCAATGGATTGGCTCCAAACCAGAC | 2277 |
| Oy | 1138 | CCACCTCTAGTCATAGCTGAGCCSCTGAGCATCACGTAATTGGAGCCCTGGGTGAC | 1197 | Db | 2242 | GTTGTCGCTGCTGGGGAGATCTCTRACTGTGTCATTGGATTGGCTCCAAACRAAT | 2301 |
| Db | 1162 | CCACCTCTAGTCATAGCTGAGCCSCTGAGCATCACGTAATTGGATTGGCTCCAAACRAAT | 1221 | Oy | 2278 | GACCTCCCGATCTCACACAGCTGAGCATCACGTAATTGGATTGGCTCCAAACRAAT | 2337 |
| Oy | 1198 | TCTCTCACGGCGCCAACTGGCGGGTCAACCTGGAAAGCTCTTGAGCTCTGACT | 1257 | Db | 2302 | GACCTCGATGTCACCCAGCTGAGCATCACGTAATTGGATTGGCTCCAAACRAAT | 2361 |
| Db | 1222 | TCTCTCACGGCGCCAACTGGCGGGTCAACCTGGAAAGCTCTTGAGCTCTGACT | 1281 | Oy | 2338 | TCTCTGGAAATGTGACCCACCTTAATATCTGGGATTAAAGAAGACCTCACA | 2397 |
| Oy | 1258 | CAGTACTGGGGCTGCTGGAGCTGGGAGAGATGAGAACATCGACACCTTACACC | 1317 | Db | 2362 | TCTCTGACAACTGACCCACCTTACCTGAGGCTACAGACAGGCTTACACGACCTCATG | 2421 |
| Db | 1282 | CAGTACCGGGCTCTCTGGAGCTGGGCGATCAGAACATCACGACGGTCACCC | 1341 | Oy | 2398 | GCCTACCCGGTGGCAGGGTATGCGCATGACAGGATGTCATGGCTTCAAGCTGTT | 2457 |
| Oy | 1318 | CTCGCGACATCTCCSGAAATCAACCTTCTGAGGGCTCTGTCGACTGG | 1377 | Db | 2422 | GCCTCGCGTGGCACTGTCGATGGCTGCACCGACCTTCACAGGCTT | 2481 |
| Db | 1342 | CTCGCCACATCTCCGGGAATCAACCTTCTGAGGCTACCC | 1401 | Oy | 2458 | CCGGACCAAAGCTGAGGATTTAGAGCTGAGGAAATGAGAGGATACAGGATAATTI | 2517 |
| Oy | 1378 | AAAAGAACCGAGCTTAATGCCCTTAAACCGGCTGGCGAGCCGACTGAGGAT | 1437 | Db | 2482 | CCGGACCAAAGCTGAGGATTTAGAGCTGAGGAAATGAGAGGATACAGGATAATTI | 2541 |
| Db | 1402 | AAGAGAGACCTTCCAAAGCTTCTGAGGACTGAGGACTGGGGAGCTGAGG | 1461 | Oy | 2518 | GATGATCATAGTAATTCTGAGGACTGAGGAAATGAGAGGATACACTT | 2577 |
| Oy | 1438 | CCTACCTGGCAGCAGGGCTGGTGGACCTGAGTGAATGAGAACATCGAGGATACACTT | 1497 | Db | 2542 | GACGACCATAGTAATTCCAGAATTCAGAAGATGAGGATACAGGATAATTI | 2601 |
| Db | 1462 | TTHATCCCCAAGCCCAACGGCTGGTGGACCTGAGAATGAGAACATCGAGGATAATTI | 1521 | Oy | 2578 | AGCGATTATGAGCTACTGAGCATGAGGATTTAGAGCTGAGGAACTCTGAG | 2637 |
| Oy | 1498 | CTAGGAAGACTGGAGATAATAACCTGTTATAGGECGCAATGACCTGTTGTC | 1557 | Db | 2602 | AGTGACTIONTGTGACTACTGAGGATTCAGGATCTGAGCTGAGGATCTGAGG | 2661 |
| Db | 1522 | CTAGGAGGATGGAGATAATAACCGTGTATGGAGCAACGACTCTGTCATTC | 1581 | Oy | 2638 | CATCTCCGATGCTTGGACAGGATCTTGTGAGGAACTCTGAGGAACTCTGAGG | 2697 |
| Oy | 1558 | AATGATCTGGTCACTATCTCCAGACTCTACAGAACATTTGGAAGGCCCTGAC | 1617 | Db | 2662 | CATCTCCGAGAGCCCTGAGGCTCCACAGAACATCGACGCCCTGAC | 2721 |
| Db | 1582 | ATGACCGGTTGGCTACTCCGCCAGACTCTACGACAAATCGACGCCCTGAC | 1641 | Oy | 2698 | GTTGACCTTCTGACCCACTATCTGGGAGGTTCTGGGAACACAGAACAGTG | 2757 |
| Oy | 1618 | ATCTCTCCATGCTAGGTCTCGGCAATTGAGCTGAGGTTCTGAGGTTCTGAC | 1677 | Db | 2722 | GTGGACTCTATGACCCAAAGCTCACGAGATGAGGATCTGAGGATCTGAGG | 2781 |
| Db | 1642 | ATCTCTCCAGCTGAGATCTCCCGGCCCTTGTGAGCTGAGGTTCTGAGGTTCTGAC | 1701 | Oy | 2758 | CCAGTGGCAGGCCAGGTTTGTGAGCTACTGCGTCTGACCTGCGGAGAACTCCAA | 2817 |
| Oy | 1678 | AACTCTGGAGACTGTGAGGAGAAAAGCTACTGCCCAGGATGTCCTCAGT | 1737 | Db | 2782 | CGATGGCAGGCCAGGCCCTGTGAGCTACTGCGTCTGAGGAGGAACTCTGAGG | 2841 |
| Db | 1702 | AACTGAGAGACGCTGTGAGGAGACTAAAGTAGCTGCGCCAGGATCTCAGGCT | 1761 | Oy | 2818 | GAGCTGCCAGCTGCTTGGAGGACTCTGAGGAGGTTCTGGGAGCATGCGAGCTG | 2877 |
| Oy | 1738 | CTGTGTCCTGTGCTCTGAGGAGAAAAGCTACTGCCCAGGATGTCCTCAGT | 1797 | Db | 2842 | GAGCTGCCAGCTGGAGGCCCTGCGACAACTGCTTGGAGGAGGAGCTG | 2901 |
| Db | 1762 | CTGTGTCCTGTGCTCTGAGGAGACTAAAGTAGCTGCGCCAGGATCTCAGT | 1821 | Oy | 2878 | GGTAGCCGGCTATGACCGAGGACTCTGAGGAGGACTCTGAGGAGGAGCTG | 2937 |
| Oy | 1798 | TCAACAGAGAGTTCAAGGAGAACCCACTGAGTGAAGTGGGGAGPATGACAA | 1857 | Db | 2902 | GAGTCGGCGCTTACGACCCGGAGGACTCTGCGAGCTG | 2961 |
| Db | 1822 | AGATCAGGAGATATCAAGGAGAACCCACTGAGTGAAGGAGGAGCTGAGGAG | 1881 | Oy | 2938 | AACATCCAGCTCTGCTGGGATGGCTCCAGATAAGTCTCTTGGCCCGAGAC | 2997 |
| Oy | 1858 | AGGAGAAATTCTACTGCTGTTGCAACGCTCTTGAAGAAGTGGACATGCCAAAGAC | 1917 | Db | 2962 | AGCATGGCTCCGCTCTGAGGAGGCCCTGCGACAACTGCTTGGCCCGAGAC | 3021 |
| Db | 1882 | ACGGACGACTCTGAGGAGTTGAGGAGACTAAAGTAGCTGCGCCAGGATCTCAGG | 1941 | Oy | 2998 | TGCACTCCACCAAACTGAGAAATTCCTCACTGAGCTGCGAGCCCTTGAGCAATATG | 3057 |
| Oy | 1918 | TCCGAGGGATGCTGTGACACTCTCTCTGCTGACTGTGTTCACTCTGGAGAG | 1977 | Db | 3022 | TGGTCCACCCAACTGAGGAGCTCTCCAGCTCTGGCAACATGCTTGGGAAACATG | 3081 |
| Db | 1942 | CGGAGACGGCTTCCGACAACTCTTCTGCGCCAGCTCTGAGGAG | 2001 | Oy | 3058 | CTTGAACACTGGAAAGAACAGAGACCTGAGCTGAGGAGGAGATGCCCATCAC | 3117 |
| Oy | 1978 | TCTCACTCCGGCAGCAGCTCTCTGAGAACATAATGCTGAGGAGCTGCGAG | 2037 | Db | 3082 | CTGGACCCCTGGAGGAAAGCTGAGCTGAGCTGAGGAGCCTCACCC | 3141 |
| Db | 2002 | GCCACGGGCAACGAGCTGAGCCCTCTGGAGAACATGCTGAGGAG | 2061 | Oy | 3118 | TCTCCACTCTGAGAACCTCTGGAAAGCTACTACAGTACAGTACAGTACAGT | 3177 |
| Oy | 2038 | AGACTGCTGATAGTGTGAAACAAGATCATCACATGCGAACCTCCAGCG | 2097 | Db | 3142 | TGCGACCCCTGGAGGAAAGCTGAGCTGAGGAGCCTCACCC | 3201 |
| Db | 2062 | ACMACACACAACTGATTTGAGGAGCACTCACATGCTGCAAAACAGGCTGAG | 2121 | Oy | 3178 | ATCAASCCAGCATGAGACTGCGAGCTCTGCTACAGAGTGAAGGTTT | 3237 |
| Oy | 2098 | TCTCTGAGGACCTACAGAACCCATGGGACCTGGCTCCATGCGAGG | 2157 | Db | 3202 | AGGAGCCAGCGCTGAGAACCTGCGACCCCTGCGACCGCTGAGGAGGAGGC | 3261 |
| Db | 2122 | TCTCTGAGTACTACAGAACAGAGCTGAGGAGCTGCGCCGAGAG | 2181 | Oy | 3238 | AATAGCTCCACCTGTCACCCAGGAGCTGAGGAGCAGTACACCTGAGG | 3297 |
| Oy | 2158 | AGAGCCCTCTGGCCTGACCCACTCTGAGGAGCTGAGGAGCAGACATCAA | 2217 | Db | 3262 | AGGGTCTCCAACTGAGCTGAGGAGATCAAGGAGATGAGTGTG | 3321 |
| Db | 2182 | AGATCCCGTCTCCCTCCACCCACTCTGAGGAGCAGGAGATATCCAA | 2241 | | | | |

| QY | 3238 CTGGGTGACTCTCTGACTACAGCAGTGGAGCTGACCAACAACTCCAGTCACCTACCC | 3357 | RESULT 4 |
|----|--|------|---|
| Db | 3322 CTGGGTGACTCTCTGACTACAGCAGTGGAGCTGACCAACAACTCCAGTCACCTACCC | 3381 | AF045454 AF045454 |
| QY | 3358 ACATCTTGAGGGACTCTCTGAGCATGAGCTGGAGGGAGGAACTTGAGACTCACCC | 3417 | DEFINITION Cavia porcellus phospholipase B (PLB) mRNA, linear ROD 02-JUN-1998 |
| Db | 3382 ATGTTCTGGAGGAGGCTCTCTGGAGCATGGTGGGGAGGGAGCTCACCC | 3441 | ACCESSION AF045454 |
| QY | 3418 ACTCTGCCAACATCTGAGGAAGTCAACCTTACCCCTTCTGCTCTCTACCGACC | 3477 | VERSION AF045454.1. GI:3172336 |
| Db | 3442 AACTGCCCCAACATCTGAGGAAGTCAACCTTACCGACC | 3501 | KEYWORDS SOURCE |
| QY | 3478 TGGAGGGACAGCAGGACTTAATGTCAGCAGGGCCAGAGCTGGACATGCCA | 3537 | ORGANISM Cavia porcellus (domestic guinea pig) |
| Db | 3502 TGGAGGGACAGCAGGACTTAATGTCAGCAGGGCCAGAGCTGGACATGCCA | 3561 | Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Buteraria; Rodentia; Hystricognathi; Caviidae; Cavia. |
| QY | 3538 GCCCAGGGCTGGACCTGTAGCGGAATGAAACAGCCCCACATCACCTGGAGAA | 3597 | Delgebeaudeuf, C., Glassama-Diagne,A., Nauze,M., Ragab,A., Li,R.Y., Capdeville,J., Ferrara,P., Fauvel,J. and Chap,H. |
| Db | 3562 GCCAAGCCGGGACTCTGGAGGAGTGAACACAGCCCTAGATGCCA | 3621 | Ectopic epididymal expression of guinea pig intestinal phospholipase B. Possible role in sperm maturation and activation by limited proteolytic digestion |
| QY | 3598 GACTGGAAGTGGACACTCTCATGGGTRACAGCTTGCTTGTGAGAT | 3657 | J. Biol. Chem. 273 (22), 13407-13414 (1998) |
| Db | 3622 GACTGGAAGTGGTCACTCTTGTGGGGCATGACTGAGAT | 3681 | MEDLINE 9826256 |
| QY | 3658 CGGAGCCACTTGGCCAGGAATATGTCAGCACATCCACAGGGCTGGACATCTC | 3717 | PUBLMED 9593672 |
| Db | 3682 CGGAGGCAGCTCGAGGGAGGAGTGTGGACATCCACAGGGCTGGACATCTC | 3741 | REFERENCE 2 (bases 1 to 4608) |
| QY | 3718 TCTGAGGAGCTCCAGGGCTTGTCACTGAGGGTGTGAGGTAGCTGGAGAG | 3777 | AUTHORS Delgebeaudeuf,C., Glassama-Diagne,A., Nauze,M., Ragab,A., Li,R.Y., Capdeville,J., Ferrara,P., Fauvel,J. and Chap,H. |
| Db | 3742 TAGGAGGACTCTGGAGGACTCTGGTCACTGAGGGTGTGAGGTAGCTGGAGAG | 3801 | TITLE Direct Submission |
| QY | 3778 TACCAAGGCAAGCGGGGAGATGTGCCA--TCTGGGAGCTGAGACTCACTTGC | 3834 | SUBMITTED (30-JAN-1998) INSERM unite 326, Hopital Purpan, Toulouse |
| Db | 3802 CACCAAGGACAAAGGGGGAGATGTGCACTGCTGGAGCTGAGCCACTCTGC | 3861 | 31059, France |
| QY | 3835 CTGAGACACTCGCAAGCTCCCTGAGAAGGAACTGAGACTGAACTC | 3894 | JOURNAL J. Biol. Chem. 273 (22), 13407-13414 (1998) |
| Db | 3862 TCAAAATACTCTGAGCTGAGCTGAGGAGTGAAGAAGTGAACCTC | 3921 | FEATURES source |
| QY | 3895 CAGCATGGCATCTCGAGTTCTCTACTCGCCACAAATACACAGCAGCTTGC | 3954 | 1. .4608 |
| Db | 3922 CAGGGGCTCTGGAGCTCTGAGGACTCTGAGGAGGAGCTGGAGCTGCC | 3981 | 1. .4608 |
| QY | 3955 GTTCTGGTCAAGCTTCTCCAAACACACTACCCACTGAGGAGGAGACT | 4014 | gene |
| Db | 3982 GTCCTGTGCAACCTTCTTCAGGAGACTCTCTGCTCCACTGAGGGAGGAACT | 4041 | CDS |
| QY | 4015 GACCTCACCTCTCTGGAGGACTGTTCACTCTGAGCCGSCATCGGAGATG | 4074 | /gene="PLB" |
| Db | 4042 GACCTCACCTCTCTGGAGGACTCTCTGAGGAGACTCTGAGGGAGT | 4101 | 1. .4431 |
| QY | 4075 GCAATGCACTCTGAGAACACATCTGGACACAGCTGGGGCGGAGAAGT | 4134 | /gene="PLB" |
| Db | 4102 GCAATTGCGCTCTGAGAACACATCTGGACAGTGGGCCAACAGAACACAC | 4161 | /note="similar to AdRab-B from adult rabbit intestine" |
| QY | 4135 TCACCCACAGCCAGCCAACTCAAGTGCCTCTCTGAGGCCCTACTCTACCC | 4194 | /codon_start=1 |
| Db | 4162 TCACCTACAGCCGACACACATCAAGTGCCTCTCTGAGGCCCTACTCTACCC | 4221 | /product="phospholipase B" |
| QY | 4195 CTGGGACAGCGGATTCTCCAGACGGCTGAAGAGCCCGAGGTGTCTACTGG | 4254 | /protein_id="AAC40129.1" |
| Db | 4222 CTCCGAAACAGCCGGCTACTCCCGACGAGCTGGAGGAGACCCACGTGCTACTGG | 4281 | /xref="IGI:3172337" |
| QY | 4255 GCTGTCCTGGAGGGAGGGAGGGCTTGTGGGGATCTGGGACAGTGTGCTGG | 4314 | /translation="MGLQPCVLUQGLLGGGTOIHTSGKSTLEGQWPEPTKONPFS |
| Db | 4282 GACTGCGCTGGAGAGGAAGGGCTCTCATGGCATCTGGCATGGAGGG | 4341 | FSCSPKRKGAMPSVHTPAPDILKQIAGDMETPPDGGAVNVDTSERTSEFKEWRG |
| QY | 4315 AGGTCGAGGAGGAGGGGGAGGAGTCTCAATGAGCTGGCTGCGCTG | 4366 | FGMMVVFLVSDITIHSNEPVSLPWTWRSVLPVLTQASERLVEUQKSNPKPQDF |
| Db | 4342 AGGGCATGCGTGTGGCCGGGGAGACCCCGCTGAGCAAGAGACAG | 4393 | QDWPWLUVNRSNLVCPSPABEPMNSNDKLQIYLHQLHEUPRPAVNMVKQLOIUNRLYKD |
| QY | | | EVVAMPRWHQGMLSRSPERACGSEBTSDKLVTVWMSQSLVQPGQKDPSPFSTERPRMKPS |
| Db | | | FAVVFOQPFPEVSVSEEPSSPEPSPDPTLAISWNHMKVPGQKDPSPFSTERPRMKPS |
| QY | | | QESPLVFTYNSVNTOSRLKQRORKESETETCPDKPDPSDTPSPVRLKPDIV |
| Db | | | GALGDSLUTANGAGSSRPGLDVLVYRQLSWSGADRNISVTLVNLREFENPL |
| QY | | | KGFSPTGKGANSVGAFNQVWAGADLIPQARTVUDMKNTNSINFEDWKLTIV |
| Db | | | IGGNDLCPDRSSDPVNTSPEUTDTNROADLILHARVPRATVUNVYKQLOIUNRLYKD |
| QY | | | SRVSCEPLRILKNCVLLDDDNSTELESIDINKKYQERTHOLTESGRDFTD |
| Db | | | VLPQPFKVDPKTSKGPOINTSFADCFNQKERSLNRSEBLVSEPLRSHARASALWQKMLVPLWPKQSKTDL |
| QY | | | NFSENSTDICPNOQARPYLSTWYKNGEKGWMLTERBERTSASPSPVHARLPDVWV |
| Db | | | HALGDSLUTANGAGSSRPGLDVLVYRQLSWSGADRNISVTLVNLREFENPL |
| QY | | | GYAVGFGDASNTNATNQVPGAKOELBLNSQVKTQVKMDPDRINFEHWKVLK |
| Db | | | CNDIUNHCPTDLYSSANFHNLALDTHREVPRAVNUVDPNNSMIPQFLGN |
| QY | | | PDCKPQKQASILCNVLSRBNSEBLRMDAIIKAYQSSRLEVRGSDTREDFSV |
| Db | | | DQPFNTRIPLERDPDTSFADCFNQKERSLNRSEBLVSEPLRSHARASALWQKMLVPLWPKQSKTDL |
| QY | | | LTADISLCPQPOEPRLRTDNTSYTPKATEWSDPCTEKPSVSPVTPSPHKL |
| Db | | | QPADTKVAALGDSLTTAVGRASNSDLMWSLGGDGALETHTLPNKK |
| QY | | | FNPSPFGPSNTGTTTGTGAGVNAVERARARQDPAQDLOVEMKASTENIEFMWDKLT |
| Db | | | LFTGSDLCHYCDNBNSBVEYVTOIROADLILHARVPRATVUNVYKQLOIUNRLYKD |
| QY | | | OGGCHTALLPQASLTSRPLFSSVTOEIKVTLQSDMSRUSYQEKTQREDFAV |
| Db | | | VVQPFFONTLPLDKLGSPTDFEDCLLRSERCHAEWALANMLNMLWPKQSKTDL |
| QY | | | NFTYNTKLIKPSTESPVLQTSLSLPPVTEKASGVAPGIVSAAAGGLVGLV |
| Db | | | ILAVSIVSSRFRQKSPSPSPVPAVNE |

BASE COUNT

ORIGIN

67.7%

Score 2962;

DB 10;

Length 4608;

Pred. No. 0;

Best Local Similarity 81.4%;

Query Match

| | | | | | | | |
|----|------|--|------|----|------|---|------|
| QY | 2155 | GACAGAGGCCCTCTGCCTTGACCCACTCTAGTCATGCCCTGAGACCTSCAGACITC | 2214 | Db | 3271 | | 3330 |
| Db | 2191 | GAGAGGACCCCTCTGCCTTGACCCACTCTAGTCATGCCCTGAGACCTSCAGACITC | 2250 | QY | 3225 | GCCTGGGTGACTCTCTGACTACAGCAGTGGAGCTGGACAAACAACTCCAGTGACTTA | 3354 |
| QY | 2215 | CAAGTTGCGCTCTCTGGGGATTCTCTGACAGTGGCTGAGACCTSCAGACITC | 2274 | Db | 3331 | GCCTGGGTGACTCTCTGACTACAGCAGTGGAGCTGGACAAACAACTCCAGTGACTTG | 3390 |
| Db | 2251 | CCTGTTGCGCTCTGGGGATTCTCTGACAGTGGCTGAGACCTSCAGACITC | 2310 | QY | 3355 | CCACATTTGGAGGGACTCTTGGGATGGAGCTGGAGACTTGAGACITC | 3414 |
| QY | 2275 | GAGACCTCCCAGTGTACCAACAGAAGCTGGACTCTCATCACAGTCAGAGTC | 2334 | Db | 3391 | CTCATGTCCTGGGGGACTCTCTGGAGCATGGAGATGGAGCTTAGAGCCAC | 3450 |
| Db | 2311 | GGGACCTGTGTGATGTGATCACACAGTATGGGGACTCTCATAGTCAGAGAC | 2370 | QY | 3415 | ACCACACTCCACATCTGAGAAGTCAACCTTACCTCTGGTTCTTACAGC | 3474 |
| QY | 2335 | GCCTCCCTGGAGATGTGACCACTTACCTATACCTGGGGATTACAGAACTC | 2394 | Db | 3451 | ACCAACTCCACATCTGAGAAGTCAACCTTACCTCTGGTTCTTACAGC | 3510 |
| Db | 2371 | GGCHCCCTGATGTGATGTGACCACTTACCTATACCTGGGGATTACAGAACTC | 2430 | QY | 3475 | ACCTGGGGGAGCAGCGGACTTAATGGGAGCTGAGACTCTGGAGCTAGGACTG | 3534 |
| QY | 2395 | AACAGCTAGCCGCTGGGAAGGGTGTGAGCCTACATCACAGGATGATCTCA | 2454 | Db | 3511 | ACCTGGAGAGCGGGGATCACACCTGAGCTGGAGGAGCCAGGCTAGGACTG | 3570 |
| Db | 2431 | ACTGGTAGGGCTGGCACTGTGATGCGAACACCAATGATCTCAACAGGCC | 2490 | QY | 3535 | CCAGCCAGGCCCTGGGACTCTGGTAGAGGATGAAACAGCCCGAGCATACCTGG | 3594 |
| QY | 2455 | GTGCCCCGAGAAAGGCTGAGATCTATGAGCAACTCTGAGAGATG | 2514 | Db | 3571 | CCAGCCAGGCCGAGACTCTGGTAGAGGATCATACCTGGCTGATTG | 3630 |
| Db | 2491 | GTTCTCTGGAGGCTGAGAAGTATGAGCAACTATGAGCAACTCTGGAGAGTG | 2550 | QY | 3595 | AAAGACTCTGAGAGCGGGGATCACACCTGAGCTGGAGGAGGGCCAGGCTAGGACTG | 3654 |
| QY | 2515 | AAAGATGATGATGAGTAATTTCATGAGACTGAGCTGAGGCTCATACGTC | 2574 | Db | 3631 | ATGAGCTGAACTGATCACCTCTCTGATGGAGCAAATGACTGTCATACTG | 3690 |
| Db | 2551 | AAGAGATGATGAGCTACATAATAATTTCACAGAGACTGAGCAACTGAGAGTG | 2610 | QY | 3655 | AATCCGGGGGCCACTTGCCAGGAAATGTTGAGCTGAGGAGCTGAGCTGAG | 3714 |
| QY | 2575 | GCGCGCGATTATGACTACTGACAGATCTGAGATCTTATGAGCCAACTGTT | 2634 | Db | 3691 | AATCCGGGAGAACCTCTGAGGAGGATCACACCTGAGAGCCAGGAGCCAGGACTG | 3750 |
| Db | 2611 | ACCAATGATGTTGTCACACTGACAGATTGGACCTGATTTGAGCAACTTTTC | 2670 | QY | 3715 | CTCTCTGAGGAGCTCCAGGGTTTCAACGTTGAGGAA -- GGTCTATGGCTCT | 3771 |
| QY | 2635 | GACCATCTCGGAAATGCGCTTGAGCTCTGAGAGGTTGCCAGAGTCTCTGTCAC | 2694 | Db | 3751 | CTGATGAGGAGCTCCAGGGCCTCTCATCAACGTAATCATCATGAGCTGCT | 3810 |
| Db | 2671 | AACCATCTCTAACGCTCTGACATCCCTACATAGAGGTTCCAGAGCCCTTGTCAC | 2730 | QY | 3772 | ACCGTGTACGAGCCAAAGGGGAAATG -- TGCATGCTGGAGCTGAGCTGAG | 3828 |
| QY | 2695 | CTCTGGACTCTCTGAACCCACTATCATGGGAGGTTCTGGGAACCCAGAACAG | 2754 | Db | 3811 | GGCTGCAACGGCCAAAGGGGAACTGAGCTGAGGAGCTGAGGAGCTG | 3870 |
| Db | 2731 | CTCGTGTGACCTCATGACCCAGTATCATGGGAGGTTCTGGAGAACCCAGACAG | 2790 | QY | 3829 | ACTGGCTGAGACTCTGAAAGCTCTGGAGAGCAGAAGTGAAGAAGTGAAGT | 3888 |
| QY | 2755 | TGCCCACTGGGAGGAGGAGGTTGGTACTCTGGGTTCTGACCTGGGGAGACTCC | 2814 | Db | 3871 | AGTGGCTCTGAGACTCTCCAGCTCTGAGGAGCTGAGGAGGACTCTGAGCTG | 3930 |
| Db | 2791 | TGCCCTGTGAGGCCAGATTGTCACACTGCTGAGCTGGAGAGACTCC | 2850 | QY | 3889 | AACCTCCGATGCTCCAGTTCTCTACTCTGGCCAACTACACAGGGAGGAC | 3948 |
| QY | 2815 | CAAGAGCTGACCGAGCTGGAGGCTCTCGGCTACCGAGCAGTCGGAGCTG | 2874 | Db | 3931 | AACCTCCGAGCTGCTCCAGGGCCTCTCATCAACGTAATCATCATGAGCTG | 3990 |
| Db | 2851 | TATGAGCTGCTGAGGAGGCTCTGACCCGAGCTACCAAGAGCAGCTGAGCTG | 2910 | QY | 3949 | TTGGGGTTGGTGTGAGCCGCTTCTCTGAAACAAACAACTACCCACTGAGGAGGG | 4008 |
| QY | 2875 | GTGGGGTCAAGGCSCTATGACACCGAGGAGCTCTGTTGAGCTGAGCTG | 2934 | Db | 3991 | TTGGGGTTGGTGTGAGCCGCTTCTCTGAAACAGCCGCTACCCGAGCTG | 4050 |
| Db | 2911 | GTGGATCATGAGGCSCTATGACACCGGGAGGACTCTGTTGAGCCCTCTG | 2970 | QY | 4009 | GACACTGACTCTCTCTGAGGAGCTGTTCTCTGAGGAGCTGAGGAGGG | 4068 |
| QY | 2935 | OAGACATCCAGCTCTGCTCTGGGATGGCTCCAGATAGTCCTCTTGGCCA | 2994 | Db | 4051 | AGCAGTGAACCTCTCTGAGGAGCTGTTCTGAGGAGCTGAGGAGGG | 4110 |
| Db | 2971 | CTCAACATCGGCTTCACTCTGGAGGTGGGGTCCAGAGACCTCCCTCTGGCCG | 3030 | QY | 4069 | GAGATGCCATCGACTCTGGACAAACTGCTGAGACAGTGGCCGAGACTCTC | 4128 |
| QY | 2995 | GACTGTCATCCACCCAAATCGGAAATTCCTCCAGCTGCCAGAGCCCTTGTGAC | 3054 | Db | 4111 | GAGATGCCATCGACTCTGGACAAACTGCTGAGACAGTGGCCGAGACTCTC | 4170 |
| Db | 3031 | GACTGTCATCCACCCAGGCCAGAAATTCCTCCACACTCTCTAGAGCCCTTGTG | 3090 | QY | 4129 | AACACTTACCCACAGCGAGCCAAATCAAGTGGCCCTCTCTGAGAGCCCTTACTC | 4188 |
| QY | 3055 | ATGCTGAAACCACTGGAGGAAACAGAGCCCTGAGCTGAGAGCTGCCATC | 3114 | Db | 4171 | AACACTTACCTACACGAGCTGAGGAGCTGAGGAGCTGAGGAGCTG | 4230 |
| Db | 3091 | ATGCTGAAACCTGTGAGGAGCAACTGAGCTGAGGAGCTGAGGAGCTGAG | 3150 | QY | 4189 | TACACCTGGAGACAGCGAGCTCTGGAGAGGAGCCGGAGGCTG | 4248 |
| QY | 3115 | ACCTGTCCTCTGAGATGCGCTCTGGAGACCCCTGGAGATGAGTAACTACAGT | 3174 | Db | 4231 | TACACCTGGAGACAGCGAGCTCTGGAGAGCTGAGGAGCTGAGGAGCTG | 4290 |
| Db | 3151 | CCCTGGCCACCCAGGAGGGCTCTGGAGAACCCCCAGAACAGGAGCTACAGT | 3210 | QY | 4249 | TACGGGGCTCCAGTGGAGGGAGGGCTCTGGAGAGCTGAGGAGCTGAGGAGCT | 4300 |
| QY | 3175 | CCCATCAAGGAGCCATGTGAGGAGCTGGGGAGTACTCTCTGTTGAGAGGAGCT | 3234 | Db | 4291 | CCCGGGATGTTGAGGAGCACTGGGCTCTGGAGAGCTGAGGAGCTGAGGAGCT | 4342 |
| Db | 3211 | CCCAACCAAGGAGCCATGTGAGGAGCTGGGGAGTACTCTGAGGAGCTGAGGAGCT | 3270 | QY | 3235 | TCCAAATAGTGTCCAACTCTGTCACCCAGGAGACATCAAGTGGTCA | 3324 |

| | | | |
|----|------|--|------|
| QY | 1143 | ACAAATTCACCTGCTCTGACAAAGACCCCTCGACATCCGCCACAGTTCACGG | 1202 |
| Db | 1156 | CTGAAAGCGGCTGACATCAACGTAATTGGAGCCCTGGGTACTCTCACGGAGCAT | 1215 |
| Db | 1203 | CTGAGGCGGCTGACATCAAGGTCTCGAGCCATGGTACTCGTCAGGAGGCAC | 1262 |
| QY | 1216 | GSGGCCTGGTCCACAACCTGGGAATGAGCTTGACTCTGACTACGAGGCTGRC | 1275 |
| Db | 1263 | GGGGCAGGGTCCAGCCCTGGGATGTCCTGGAGATCTAACGAGGCTGRC | 1322 |
| QY | 1276 | TGGAGCTGGGAGAAGTGGAGAACATCGCACCGTTACACCTGGCRACTCTCGG | 1335 |
| Db | 1443 | GCATCTCAACCPGGCGTAGGAGGAGCCAAATCTGATGGTTAGCTGCCAGGAA | 1502 |
| QY | 1456 | AGCTCTGGGACTCTGATGAGATGACAGGAGTACCTTCAGGAGACTGGAGATA | 1515 |
| Db | 1503 | AAGCTGGTGGAGCTGATGAGGATGAGGATGAGGAGACTGGAGATA | 1562 |
| QY | 1516 | ATRACCTGTTAGGGCAATGACCTCTGTCATCTGCAATGATCTGGTCACTAT | 1575 |
| Db | 1563 | ATCAGCTGGTTAGGGCAACGACTCTGGCTCTGCATAAATCTGGTGCRT | 1622 |
| QY | 1576 | TCTCCCCCAGAACCTTCACAGACAACTATGGAAAGGCCCTGACATCTCATGTCAGGTT | 1635 |
| Db | 1623 | TCTCCCCAACCTTCACAGACACATCAAGACGCCCTGACATCTCATGTCAGGTT | 1682 |
| QY | 1636 | CCTGGGCAATTGGAACCTGGTGGAGCTGGTGTGACATCTGGTGTAC | 1695 |
| Db | 1683 | CCCGGGGCACTTGTGACATGGTCTGGTGTGACATCTGGTGTAC | 1742 |
| QY | 1696 | CAGGAGAAAAGTCTACTGCCAGGATGATCTCAGGCTCTGTGTCCTGTCCTG | 1755 |
| Db | 1743 | AATGAACTTAAGTCAGCTGGCCACGGATGATCTCAGGCGTGTGTCCTGTC | 1802 |
| QY | 1756 | AAAGTTGATGATACATCAACAGAACCTCTACCCATCGAAATCGAAAGTTTCAG | 1815 |
| Db | 1803 | AACCTGGTGGAGACTTCACCGAGAACTGGCCAATCTGGAGATCTGGTCAAGAAGGAGTATCG | 1862 |
| QY | 1816 | GAGAAGACCCACCACTGGAGATGGGGAATGACACAGGAGATTTCAG | 1875 |
| Db | 1863 | GGAGAAACTGGAAACATGATGGAGATGGGGAATGACACAGGAGACTTCACCG | 1922 |
| QY | 1876 | GTGTCGACCGGTTCTGAAACGGACATGCCAACGACTCGGAAGGATTCGAC | 1935 |
| Db | 1923 | GTCTCCAGCCATGTTGAATGTCATGCCACGACCTGAGGGCTGCCAC | 1982 |
| QY | 1936 | AACTCTTCTGCTCTCTGACTGTTCCACTTCAGGAGAACGTTACTCCAGGCC | 1995 |
| Db | 1983 | AGCTCTTCTGCTGAGAACATATGCTGGAGCTGTTCCACTTCAGGAGAACGAGC | 2042 |
| QY | 1996 | AGTGCTCTGCTGAGAACATATGCTGGAGCTGTTCCACTTCAGGAGAACGAGC | 2055 |
| Db | 2043 | ATGCCCTCTGGAGAACATGCTGGAGCTGTTCCACTTCAGGAGAACGAGC | 2102 |
| QY | 2056 | GAAGAACGATCATACATGTCGACCGGTTCTGAGGACTACAG | 2115 |
| Db | 2103 | GAATCAAGTCCTATGCTGTCACACAGACCTACCGGTTCTGAGCACCAAG | 2162 |
| QY | 2116 | AACAGCAGTCAGGTCTGGGACCTGGCTGAGGAGAACGAGCCTTCGCT | 2175 |
| Db | 2163 | AACAGCAACTGGGACATGGAACTCTGATGTCATGAGGAGAACCCCTGCTCA | 2222 |
| QY | 2176 | CACCTCTCTGCTGAGACCTGCTGAGAACATCAAGTGTGGGCGCCCTGGG | 2235 |
| Db | 2223 | CCACCAACTCTGCTGAGAACCTGAGACCTGCTGAGAACATCAAGTGTGGGAGCTGTCAGGAA | 2282 |
| QY | 2235 | GATTCTCTGACCGCTGGCATGGATTGGCTCAAACCGAGACACTCCGGAGTCACC | 2295 |
| Db | 2283 | GACTCTGTAATGTCAGGCTCCAGAAGTGAATGACTCTCCAGAAGAGTGATCT | 2342 |
| QY | 2296 | ACAGATAGGGACTGTCATACAGTGGAGGGAGCGCTCTGGAGATGTGACC | 2355 |
| Db | 2343 | ACACGATATGAGACTGTCATACAGTGGAGGGAGCGCTCTGGAGATGTGACC | 2402 |
| QY | 2356 | ACCTTACCTTAATCTCGGAGTTAACAGAACCTACAGGCTACGGCGACG | 2415 |
| Db | 2403 | ACCTGCCACATCTCGGGAAATTAAATGAAATCTGAGACAGGACTACTCGTGGAA | 2462 |
| QY | 2416 | GGTGTGCAAGTCAGTCAGACTGTCATGAGCTGGTGTCCGGAGCAAGCTGAG | 2475 |
| Db | 2463 | GTTGAGCTACTCTGCAAGGGCTCCCTAACAGGCTGTTCTGGGAGAAAGCTGAG | 2522 |
| QY | 2476 | GATCTTATGAGGCAAGTCAGTCAGACTCTGATGTCAGAGATGAGATGATAGTAAT | 2535 |
| Db | 2523 | ACACUTGCAAGTCAGTCAGACTCTGATGTCAGAGATGAGATGACACAGAGTCAC | 2582 |
| QY | 2536 | TCCATGAGACTGGAGGACTCATCGAGAGATGAGATGAGATGAGAATGACACAGAGTCAGGAGA | 2595 |
| Db | 2583 | TTTACCAAGACTGGAGGTCATCTGATGATGATGTTGGGGCAGGGACTGTGTCAC | 2642 |
| QY | 2596 | TGACAGAGTTGAGATCTGATCTGCACTTCCTGACACTTGTGAGACATCTGGGAGCAAGCTGAG | 2655 |
| Db | 2643 | TGCAAGGATTCAGTGAATGTCAGCTTCCTAACAGGCTGTTCTGGGAGAAAGCTGAG | 2702 |
| QY | 2656 | GACGCTCTGATAGAGGAGTGGTCACTCGAGGAGCTCTGGTCAACTCTGGTCAACCTCTGGACTCTCTGAA | 2715 |
| Db | 2703 | GACATCCCTCATTAAGGGAGTACCCAGAGSACCTCTGCAACTCTGGACTCTCTGAA | 2762 |
| QY | 2716 | ACTATCATCGGAGGTGTCCTGGAGCTGGTGTGACATCCAGGAGCTGTC | 2775 |
| Db | 2763 | AGTATCATCGCAAGTGTCTGAGAAGACCCAGACAGTCGGCTGTCAGGAGT | 2822 |
| QY | 2776 | GTTRGTGTAATCTGGCTGACCTCTGGGGAGACTCCAGAGCTAACGGCTGGAG | 2835 |
| Db | 2823 | GTCCCTGTCACATCTGGTCTGGACCCAGGGAGATTCCTCATGACTGGAGGTGGAG | 2882 |
| QY | 2836 | GCCTTCAGCGAGGCTACCGAGGAGCATGGCGAGCTGGTGTGGGTCAAGCCGCTATGAC | 2895 |
| Db | 2883 | GCCTTCACCAAATCTTACCAAGAGTAGAGCTGGAGCTTCACGTTGTCAGGAGCAGC | 2942 |
| QY | 2896 | ACGGAGGAGGACTCTCTGAGCTGGTGTGAGCTGGAGCCCTCTCCAGAACATCCAGCTGTC | 2955 |
| Db | 2943 | ACCCGGAGGAGTCTCTGAGTACTGAGCTGGAGCCCTTCCTCACATCAGGCTCCAT | 3002 |
| QY | 2956 | CTGGGGATGGCTCCAGATACCTCTTGGCCAGATGGCATCCACCAATACAG | 3015 |
| Db | 3003 | CTAGAGAATGGAAATCCAGATACATCTCTTGGCCAGATGATCCCTCTAGGCCAG | 3062 |
| QY | 3016 | AAATTCCACCTCCAGCTGCCAGAGCCCTTGGCCANTATGCTGAACTCTGAG | 3075 |
| Db | 3063 | AACTCCACACTCTGCTCCGGAGGCCCTTGGCCANTATGCTGAACTCTGAG | 3122 |
| QY | 3076 | AAACAGAGACCCCTGGAGCTGGAGAGGACTCTGAGCTGGAGGCTTCAGATGAG | 3135 |
| Db | 3123 | AAATGGATACCTGGACCGAGAACCTCATAGCTTGGCTGCCACCAAGGACAG | 3182 |
| QY | 3135 | CCCTCTGAGAACCCCTGGAGTAGTACTACGAGCTGCCATCAASCGCATGG | 3195 |
| Db | 3183 | CCCTCTGAGAACCTCTGGAGACTACAGTACCTTCAAGCCAGCATGCT | 3242 |
| QY | 3196 | ACTGGGGAGTGAATCTCTGGTACAGAGTGGAGGCTTCAGATGTTCAACCT | 3255 |
| Db | 3243 | ATTTGGGAGTGAATCTCTGGTACAGAGTGGAGGCTTCAGATGTTCAACCT | 3302 |
| QY | 3256 | GTCCACCACTCCGACAGACATCAAGTGTGGGCGCCCTGGTGTACTCTGTACT | 3315 |
| Db | 3303 | GTTCATGAGCTCCGACCATCAGACATCAAGTGTGGAGGAGCAATGGGACTCTTCGAC | 3362 |

| | | | | | | | |
|----|------|---|------|----|------|--|------|
| Db | 2583 | TTCACAGAACAGCTGGAAAGTCACACTGTGATGGGCCAGCAGTGTGACTTC | 2642 | Oy | 3676 | ACCGAATTGTTCAAGCACATCCACAGCCCCCTGACATCCTCTCTGAGGAGCTCCAGG | 3735 |
| Oy | 2596 | TTCACAGAACAGCTGGAAAGTCACACTGTGATGGGCCAGCAGTGTGACTTC | 2655 | Db | 3723 | AGGAAGTGTCAAGTACATCCAGCATCCCTTGAGACTCTCTATGAGGAGTTCCGG | 3782 |
| Db | 2643 | TGCRAGGATTGCGATCTGTTACTCTGCACTGAGCCATTTCATGCATCTCGCATGCCTG | 2702 | Oy | 3736 | GCTTCGCAAGGGGGAGTCATGGAGCTGGCTGAGCTTACCGGGCAAGGGGG | 3795 |
| Oy | 2656 | GCGTCCTGCGATAGAGCTGGCCAGACTCTCTGCTAACCTGTGACTTCCTGAA | 2715 | Db | 3783 | GTTTCATCAACTGGTGAAGTCATGGAGCTGGCTCCGGCTCTCCAGACCGGGCGG | 3842 |
| Db | 2703 | GACATCCCTGCAATTAGGGAGTACCCAGACGCCCTGTCACTTGTGACTTCATGAACCC | 2762 | Oy | 3796 | AAATGTCGCGATC--TGGCAGCTCAGAACACTGCGACTTGCCTCAGAACACTGCAANG | 3852 |
| Oy | 2716 | ACTATCATGCGGAGGTTCTGGAAACCCAGACAGTGGCGAGTGGAGGCGC | 2775 | Db | 3843 | AAATGTCGCGATCAGGGAGATTCATAGCTGCAAGTGGAGTGGAGTAAAGCTCAAAAC | 3902 |
| Db | 2763 | AATGATCATTCGCGAGGTTCTGGAAAGAACCCAGACAGTGGCGAGTGGAGTGGAGT | 2822 | Oy | 3853 | TCCTGGAGAACAGACTGAGAAAGTGAACACTGGACCTCCAGCATGGCATCTCGAGT | 3912 |
| Oy | 2776 | GTGTTGTTAACTGGCGTCTGACCGCTGGGGAACTCCAAAGCTAGCCAGGGCTGG | 2835 | Db | 3903 | CCTATGGCGATCAGGGAGATTCATAGCTGCAAGTGGAGTGGAGTGGAGTGGAG | 3962 |
| Db | 2823 | GTCTCTGTCACGTGTCTGACCCAGGGAGGATTCATAGCTGCAAGTGGAGTGGAG | 2882 | Oy | 3913 | TCCTGGAGAACAGACTGAGAAAGTGAACACTGGACCTTCAGGCTTCAGGCTTC | 3972 |
| Oy | 2836 | GCCTTCAGCCGAGCCTACCGGCGAGCTGGGGTGGGGTGGGGCTATGAC | 2895 | Db | 3963 | CTCTCTACTGGCATCGTACATGAGCTGAGCTGGCTCAGCTGAGCTTCAGGCTTC | 4022 |
| Db | 2883 | GCCTTCACCAAATCTTACCAAGACTAGATGCTCACTGGTAGTGGAGCTGGCTACGAC | 2942 | Oy | 3973 | TTCACAAACACACTCACCCACTGAGCGAGAGGGGACTGACCTCACCTCTCTCC | 4032 |
| Oy | 2896 | ACCGAGGAGGACTCTCTGTGCTGCTGCTGAGCTGACCCCTCTCACACATCCAGGCTCTGTC | 2955 | Db | 4023 | TCCTGGATACCTTGTCCCACCTGGATGAGCTGGAGCTGGCTGACCTTCTCTCT | 4082 |
| Db | 2943 | ACCGGGGGAGGTTCTCTGTGACTGAGCCCTTCCTCACACATGGCTCCAT | 3002 | Oy | 4033 | GAGGACTGTTTCACTTCAGACCGGGGAGATGCCATGCACTCTGGAC | 4092 |
| Oy | 2956 | CTCGCGGATGGCTCCAGATACTGCTCTGGCCAGATGAGCTCCACCCAATCAG | 3015 | Db | 4083 | GAAGACTTCTTCACTTCAGTCGGTGGCATGCTGAGTGGCATCTGGCTTCAG | 4142 |
| Db | 3003 | CTAGAGAAATGGGAAATCCATACCTTGTGCGGAGATTCAGTGGCATCTGGCTAACCC | 3062 | Oy | 4093 | ACATGCTGGACCACTGGGGCGCAAGACTACTTCACCCACAACTTCAACCAAGGCC | 4152 |
| Oy | 3016 | AAATTCCACTCCAGCTGGCGAGCCATTGACCAATATCTGACCACTTGGAGC | 3075 | Db | 4143 | ACATGCTGGACCACTGGCGAGACAACTTCATACACCGAAC | 4202 |
| Db | 3063 | AAGTTCACACTCACTGCGAGAGCCCTTGGCCATATCTGACCCCTGGAAAG | 3122 | Oy | 4153 | AAACTCACTGCTCCCTCTGAGGCCCTACTCTCACCCCTGCGGACAGCCATG | 4212 |
| Oy | 3076 | AAACAGAGACCCCTGGAGCTGAGAGCAGATGCTCCATCACCTGCCCACACTAGAATGAG | 3135 | Db | 4203 | AAACTCACTGCTCCCTCTGAGGCCCTACTCTCACCCCTGCGGACAGCTGAGT | 4262 |
| Db | 3123 | AAATGGATACTTGTGACCGGAAAGTCACTGCTTGGCTGGCCACCCAGGAGAAG | 3182 | Oy | 4213 | CTCCCGAGGAGCTGAGAAAGCCCCGGAGGTGCTCTACTGGCTGTCCTCACGCG | 4272 |
| Oy | 3136 | CCCTCCCTGAGAACCCCTGGAGATGAGCTACGTCACCTCAAGCCACCATGGAG | 3195 | Db | 4263 | CTCTAGCAAGGCTAAGAANACTCCATACACTCTACTGGCAGTCCTGGCTCA | 4322 |
| Db | 3183 | CCCTTCCTGAGAACCTTGGAGACAGTAACAGTACCTACGTAACCTGAG | 3242 | Oy | 4223 | GGATCGCCCTGTGGGGCATCTGGGACAGTGGCTGCTGGGGCCAGAGGAGTC | 4332 |
| Oy | 3196 | ACTGGCCAGACTCTGTGCTACAGTGGAGGCTCCATAGTGTCAACCT | 3255 | Db | 4323 | GTAGTGGCCCTGTGGGGTAGTGTGGCATCTGGGAGACTGTGGAGACTCGTC | 4382 |
| Db | 3243 | AAATTGGGGCACTTCTGTGTCACAGAGCAGAGTCCTCCAGCAAGGTACCC | 3302 | Oy | 4333 | CGAGGGAGA 4343 | |
| Oy | 3256 | GTCCACCAAGCTCGACCAAGAACATCAAGTGGTGGCCGCGCTGGGTACTCTGACT | 3315 | Db | 4383 | CTATGAGGA 4393 | |
| Db | 3303 | GTCTCATGAGCTGGACCATCACAACTTGTGGAGGAGCATGGTGGACT | 3362 | | | | |
| Oy | 3316 | ATGAGCACTGGGCTCTGACCAACAACTTGTGGGGACTC | 3375 | | | | |
| Db | 3363 | ACAGCCACGGGAGCTCGACCAACTGAGTCCTCACAGTCAGACCCCTGGAGGGCTG | 3422 | | | | |
| Oy | 3376 | TCTTGGGACATCTGGGGATGGAACTGGGACTCACACACACTCCGACATCTG | 3435 | | | | |
| Db | 3423 | TCTTGGGACATCTGGGGATGGAACTGGGACTCACACACTCCGACATCTG | 3482 | | | | |
| Oy | 3436 | AGAGGTCAACCTTACCTCTGGCTCTTACCAAGCACCTGGAGGGAGCAGCAGGA | 3495 | | | | |
| Db | 3483 | AGAGGTCAACCTTCTGGATCTGGACCTGGAGGAGACGGCGAGGA | 3542 | | | | |
| Oy | 3496 | CTAAATGTCGCGGGGGAGGGGGAGCTAACGGACATCCAGCCGGCCAGCTG | 3555 | | | | |
| Db | 3543 | TAAATGTCGCGAGAACGGGGAGCTAACGGACATCCAGCCGGCCAGCTG | 3602 | | | | |
| Oy | 3556 | GTAGAGGAGATGAAAGCCGCGACATCACCTGGAGAAAGCTGGAGGGTGTACA | 3615 | | | | |
| Db | 3603 | GTGAGAGAAGATGAAAGACCCCTACAACTAACATAGAGAAGTGTGATTACA | 3662 | | | | |
| Oy | 3616 | CTCTCATGGGGTCAAGACTGTGTGATTTCTGTGAGAATCGGAGGCCACTGGCC | 3675 | | | | |
| Db | 3663 | CTCTCATGGGGTCAAGACTGTGTGATTTCTGTGAGAATCGGAGGCCACTGGCC | 3722 | | | | |

RESULT 7
AX711962
LOCUS AX711962 Sequence 1 from Patent WO20062977.
DEFINITION 1835 bp DNA linear - PAT 10-APR-2003
ACCESSION AX711962
VERSION AX711962.1 GI:29787748
KEYWORD
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1. Yan, C., Ketchum, K., di Francesco, V. and Beasley, E.M.
TITLE Human phospholipase b-like Polypeptide and uses thereof
PATENT: WO 20062977A1 15-AUG-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
Source 1..1835
Organism="Homo sapiens"
mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 458 a 549 c 474 g 354 t
ORIGIN

| | | |
|-----------------------|--|--|
| Query Match | 31.1% | Score 1362; DB 6; Length 1835; |
| Best Local Similarity | 89.3% | Pred. No. 0; Mismatches 0; Indels 190; Gaps 3; |
| Matches | 1582; Conservative | |
| OY | 2618 | CTCGCAGCCACTTGTGACCATCTCCCAATGCCCTGGACGTCGCTGATAG----- 2671 |
| Db | 1 | CTGCGAGCCACTTGTGACCATCTCCGAATGCCCTGGACGTCGCTGATAG----- 60 |
| OY | 2672 | - - - - - ACGTGCAGAGCTCTGTCAACTCTGGACTCTGGACTCTGAACTCTGCAATG----- 2725 |
| Db | 61 | TCCCTTAGCTGCCAGAAGCTGCTGAACTCTGCAACCTCTGAACTCTGCAATG----- 120 |
| OY | 2726 | GCGAGGTGTTCTGGAAACCCAGACAAGTGCCAGTGGAGGCTTGAGGTTGTGTA 2785 |
| Db | 121 | GCGAGGTGTTCTGGAAACCCAGACAAGTGCCAGTGGAGGCTTGAGGCTTGAGGCA----- 168 |
| OY | 2786 | AAGCGTGTCTGACCTCTGGGGAGAACTCCAAGAGCTGAGCTGACCTGAGCTGGCTTCAC 2845 |
| Db | 169 | - - - - - ----- 168 |
| OY | 2846 | GAGCCTAACGGAGCAGATGGGGAGCTGGGGCTATGACACCCAGGG 2905 |
| Db | 169 | - - - - - GAGCAGATGGGGAGCTGGGGCTATGACACCCAGGG 218 |
| OY | 2906 | ACTCTCTGTGCTGTGACGCCCTCTCCAGAACATCCAGCTCCCTGTCTGGGGAT 2965 |
| Db | 219 | ACTCTCTGTGCTGTGACGCCCTCTCCAGAACATCCAGCTCCCTGTCTGGG----- 273 |
| OY | 2966 | GGCTCCAGATACTGCTCTTGTGCCCCAGATCGCATCCACCAAATGAGAAATTCCACT 3025 |
| Db | 274 | - - - - - ----- 273 |
| OY | 3026 | CCCAGCTGCCAGAGCCCTTGACCAATATGCTGAAACCACTTGAGAAGAACAGGA 3085 |
| Db | 274 | - - - - - GCTGTGAAACCACTTGAGAAGAACAGGA 302 |
| OY | 3086 | CCCTGGACCTGAGGAGAGATGGCCATACCTGTGAGATAGCCCTTCCTGA 3145 |
| Db | 303 | CCCTGGACCTGAGGAGAGATGGCCATACCTGTGAGATAGCCCTTCCTGA 362 |
| OY | 3146 | GAACCCCTGGAAATAGTAACTACAGTACCCATCAAGCACCCATTGAGAATGGGCA 3205 |
| Db | 363 | GAACCCCTGGAAATAGTAACTACAGTACCCATCAAGCACCCATTGAGAATGGGCA 422 |
| OY | 3206 | GTCATCTCTGTACAGGTGAGBGCCTCAATGTTCAACCTGTGACCCACCGC 3265 |
| Db | 423 | GTCATCTCTGTACAGGTGAGBGCCTCAATGTTCAACCTGTGACCCACCGC 482 |
| OY | 3266 | TCCGACCAAGACATCAAGTGTGGCCCTGGGTGACTCTCTGACTAAGCAGTGG 3325 |
| Db | 483 | TCCGACCAAGACATCAAGTGTGGCCCTGGGTGACTCTCTGACTAAGCAGTGG 542 |
| OY | 3326 | GAGCTGACCAAGACATCAAGTGTGGCCCTGGGTGACTCTCTGACTAAGCAGTGG 3385 |
| Db | 543 | GAGCTGACCAAGACATCAAGTGTGGCCCTGGGTGACTCTCTGACTAAGCAGTGG 602 |
| OY | 3386 | TGGAGGAGATGAACTCTGGAGACTCACACACACTGCCAACATCTGAGAGACTCA 3445 |
| Db | 603 | TGGAGGAGATGAACTCTGGAGACTCACACACACTGCCAACATCTGAGAGACTCA 662 |
| OY | 3446 | ACCCCTACTCTGGCTCTTACCGAGACCTGGAGGGACAGAGCTAATGTGG 3505 |
| Db | 663 | ACCCCTACTCTGGCTCTTACCGAGACCTGGAGGGACAGAGCTAATGTGG 722 |
| OY | 3506 | CASCGGAAGGGCCAGACATGCCAACCCAGGCCCTGGACCTGGAGACTAATGTGG 3565 |
| Db | 723 | CAGGGAGAGGGCCAGAGCTGGACATGCCAACCCAGGCCCTGGACCTGGAGACTAATGTGG 782 |
| OY | 3566 | TGAAAACAGCCCCGACATCACTGGAGAAAGCTGGAGCTGGTACACTCTCTG 3625 |
| Db | 783 | TGAAAACAGCCCCGACATCACTGGAGAAAGCTGGAGCTGGTACACTCTCTG 842 |
| RESULT 8 | | |
| | BC042674 | |
| DEFINITION | Homo sapiens | 1790 bp mRNA linear . PRI 09-JAN-2003 |
| ORGANISM | Homo sapiens | Similar to phospholipase B, clone MGC:35447 |
| ACCESSION | IMAGB:5191712 | mRNA, complete cds. |
| VERSION | BC042674 | |
| KEYWORDS | MGC. | |
| SOURCE | Homo sapiens (human) | |
| ORGANISM | Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Buteria; Primates; Catarrhini; Hominidae; Homo. | |
| REFERENCE | 1 (bases 1 to 1790) | |
| AUTHORS | Strausberg, R. | |
| TITLE | Direct Submission | |
| JOURNAL | Submitted (02-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | |
| REMARK | NIH-MGC Project URL: http://mgc.nci.nih.gov | |

| | |
|---|---|
| KEYWORDS | Homo sapiens |
| SOURCE | (human) |
| ORGANISM | Homo sapiens |
| AUTHORS | Bukayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Cetarrhini; Hominidae; Homo. |
| REFERENCE | 1 Isogai,T., Sugiama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S., Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R., Tamochika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahori,K. and Matsuo,I. |
| JOURNAL | Full-length cDNAs |
| FEATURES | source |
| BASE COUNT | 691 a 597 c 591 g 598 t |
| ORIGIN | |
| Query Match | 22.2%; Score 970; DB 6; Length 2477; |
| Best Local Similarity | 99.5%; Preq. 0; Mismatches 5; Indels 0; Gaps 0; |
| Matches | 973; conservative |
| Qy | 1279 AGGCTCGGGAGATGAGACATCGGACCGTACCCCTGGCGAACATCTCGCGAA 1338 |
| Db | 6 AGCTCGGGGGAGATGAGACATCGGACCGTACCCCTGGCGAACATCTCGCGAA 65 |
| Qy | 1339 TCTAACCGGGGTGCGAGGAGGCCAGCTGAGGACTCTACTGTGCAGGCCAGGG 1398 |
| Db | 66 TCTAACCGGGTCCCTGAGACAGCTGAGGACTCTACTGTGCAGGCCAGGG 125 |
| Qy | 1399 TCTAACCGGGGTGCGAGGAGGCCAGCTGAGGACTCTACTGTGCAGGCCAGGG 1458 |
| Db | 126 TCTAACCGGGTGCAGGGCCAGCTGAGGACTCTACTGTGCAGGCCAGGG 185 |
| Qy | 1459 CTGGTGGACCTGATGAGATGACAGGAGATACACTTCAGGAGACTGGAGATAA 1518 |
| Db | 186 CTGGTGGACCTGATGAGATGACAGGAGATACACTTCAGGAGACTGGAGATAA 245 |
| Qy | 1519 ACCCTGTTTATAGCGGCCATGAGCTCTGTGATTCTGCAATGATCTGTCACATCT 1578 |
| Db | 246 ACCCTGTTTATAGCGGCCATGAGCTCTGTGATTCTGCAATGATCTGTCACATCT 305 |
| Qy | 1579 CCCAGAACTTCAAGACAGACATGGAAAGGCCCTGGACATCCTOCATGCTGGTCT 1638 |
| Db | 306 CCCAGAACTTCAAGACAGACATGGAAAGGCCCTGGACATCCTOCATGCTGGTCT 365 |
| Qy | 1639 CGGGCATTTGGACCTGGCGCTGCTGGAGATCGTCAACCTGAGGGAGTGACAG 1698 |
| Db | 366 CGGGCATTTGGACCTGGCGCTGCTGGAGATCGTCAACCTGAGGGAGTGACAG 425 |
| Qy | 1699 GAGAAAAGTCTACTGCCAAGATGATCTGGCTCTGTCCTGTCCTGAG 1758 |
| Db | 426 GAGAAAAGTCTACTGCCAAGATGATCTGGCTCTGTCCTGTCCTGAG 485 |
| Qy | 1759 TTGATGATACTCAACAGAACATGTCACATGTCACATGATCTGAG 1818 |
| Db | 486 TTGATGATACTCAACAGAACATGTCACATGATCTGAG 545 |
| Qy | 1819 AGGACCAACACTGATGAGAGTGGGAGATGACAGACAGGAGATTTACTGTGTT 1878 |
| Db | 546 AGAACCAACACTGATGAGAGTGGGAGATGACAGACAGGAGATTTACTGTGTT 605 |
| Qy | 1879 GTGCAAGCTCTTCTGAGACATGCCAAGACCTGGAGATGCTGACAC 1938 |
| Db | 606 GTGCAAGCTCTTCTGAGACATGCCAAGACCTGGAGATGCTGACAC 665 |
| Qy | 1939 TCTTCTCTGCTCTGACTGTTCACTGGCACTCTGCTGCTGCTGCTGAG 1998 |
| Db | 666 TCTTCTCTGCTCTGACTGTTCACTGGCACTCTGCTGCTGCTGCTGAG 725 |
| FEATURES | source |
| Qy | 1999 GCTCTCTGACATATCTCTGGAGCTCTTGGCCAGAGACACTCGTCATAAGTTGA 2058 |
| Do | 726 GCTCTCTGACATATCTCTGGAGCTCTTGGCCAGAGACACTCGTCATAAGTTGA 785 |
| Qy | 2059 AACAGATCAATACACTGTCGAACAGGCTTCAGGGCTTCAGGACCTACAGAC 2118 |
| Do | 786 AACAGATCAATACACTGTCGAACAGGCTTCAGGGCTTCAGGACCTACAGAC 845 |
| Qy | 2119 AGCATGCGGGTCACTGGGACCTGCTGCGCATGCGGGACAGGCCCCCTCTGCCTGAC 2178 |
| Do | 846 AGCATGCGGGTCACTGGGACCTGCTGCGCATGCGGGACAGGCCCCCTCTGCCTGAC 905 |
| Qy | 2179 CCTACCTGATGCGCTGAGACCTCAGACATCCAGTGTGCGCTCTGGGGAT 2238 |
| Do | 906 CCTACCTGATGCGCTGAGACCTCAGACATCCAGTGTGCGCTCTGGGGAT 965 |
| Qy | 2229 TCTGTGACCGTGCACAT 2256 |
| Do | 966 TCTCTGACCGTGAAGACT 983 |
| RESULT | 10 |
| AK055428 | AK055428 |
| LOCUS | AK055428 |
| DEFINITION | AK055428 |
| Homologous cDNA FLJ30866 Fis, clone FBRA020410, highly similar to phospholipase A2R-B PRECURSOR (EC 3.1.1.-). | Homologous cDNA FLJ30866 Fis, clone FBRA020410, highly similar to phospholipase A2R-B PRECURSOR (EC 3.1.1.-). |
| ACCESSION | AK055428 |
| VERSION | AK055428.1 |
| KEYWORDS | GI:16550150 |
| SOURCE | oligo capping; fib (full insert sequence). |
| ORGANISM | Homo sapiens (human) |
| AUTHORS | Bukayota; Metazoa; Primates; Cetarrhini; Hominidae; Homo. |
| REFERENCE | 1 Nishi,T., Nakagawa,S., Senoh,A., Mizuuchi,H., Inagaki,H., Isogai,T., Sugiama,T., Otsuki,T., Saito,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai,Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Watanabe,M., Murakawa,K., Kanehori,K., Takahashi,Fujii,A., Oshima,A., Sugiama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahori,K., Masuho,Y., Nagai,K. and Isogai,T. |
| TITLE | NEDO human cDNA sequencing project |
| JOURNAL | Unpublished |
| REFERENCE | 2 (bases 1 to 2477) |
| AUTHORS | Isogai,T., Otsuki,T. and Sugiama,T. |
| TITLE | |
| JOURNAL | Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 299-0812, Japan (E-mail:genomics@hri.co.jp), Tel:81-438-55-3975, Fax:81-438-55-3986) |
| COMMENT | NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RABI); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' - & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI. |
| FEATURES | location/Qualifiers |
| Qy | 1. -2477. |
| Db | /organism="Homo sapiens" |
| Qy | /mol_type="mRNA" |
| Db | /db_xref="taxon:9606" |
| Qy | /clone="FBRA20410" |
| Db | /tissue_type="Brain" |
| Qy | /clone_id="FBRA2" |
| Db | /dev_stage="fetus" |
| Qy | /note="cloning vector: pME18SFL3" |
| Db | 198. .1010 |
| Qy | /note="unnamed protein product" |
| Db | /codon_start=1 |
| Qy | /protein_id="BAH70920_1" |
| Db | /db_xref="GI:16550151" |

/transliteration="MKKNTRIHFQEDWKITLFIQGNDLDFCNLDLHHSPQNFTDNI
GRADLJHAEVPRAFVNLYVLEVNRLYOEKCYCPEMLASLCPCVLFKDNTN
ELATLIEENKKFOKTHOLIESGRYTFRDPLTYVUPPFENWMDPCKTSSLGDFEA
POCFHRFSKSHSAASALANMPLBPGQCTTRHKEFNKINTCPQVQPLRTYGM
OCHGIWIFPCRDRAPSAHLTSVHALRPAFDQIVQVVAAGLDSLTVRLGPQVVG"
ORIGIN
BASE COUNT
691 a 597 c 591 g 598 t

| | | | |
|--|--|--|--|
| | /codon_start=1 /product="hypothetical protein MGCA0917" /protein_id="AAB33606_1" /db_xref="GI:21706930" /transl_id="231108" /translation="MQLBKLMNNTQSGISLHLYRHYMEREDPAVTQPFRRNFTIP LNERCBLDLTFSDFYFSGHMAIAWNMILEPVGKTSNNFIYTRKCP SPERFLYLPSLQDKEEPSNLWAVPVALIGGLAVGILGVMLWRIVKPQOE BEEPDLPNTSVTODAVEKRULKAGN" ORIGIN | BASE COUNT 266 a 266 a 197 c 188 g 169 t | /db_xref="taxon:9606" 80 c 62 g 50 t |
| Query Match | 7.4%; Score 322.2; DB 10; Length 820; Best Local Similarity 75.1%; Pred. No. 3e-73; Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | 5.7%; Score 251; DB 6; Length 265; Best Local Similarity 100.0%; Pred. No. 1.6e-54; Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | 5.7%; Score 251; DB 6; Length 265; Best Local Similarity 100.0%; Pred. No. 1.6e-54; Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| Oy | 3809 TGGCAGCTCAGAACACTGCACTTGCCCTCAGACACTCGCAAACTCCCTGGAGAGCAG 3868 41 TGCTTGTAGGAAAACCTGCAGTCAGACTCCAAACCTCACAGATGCCAG 100 | 3042 CCTTTGACCAATTATGCTGAACACTTGGAAGCAAACAGAGACCCCTGGACCTGTGAGGC 3101 41 TGCTTGTAGGAAAACCTGCAGTCAGACTCCAAACCTCACAGATGCCAG 134 | 3102 AGAGATGCCATCACCTTCCCCTCAGAATGAGACCCCTGGAAATAG 3161 41 TGCTTGTAGGAAAACCTGCAGTCAGACTCCAAACCTCACAGATGCCAG 74 |
| Oy | 3869 ACTGAGAAAGTGAACCTGGAACTCTGGCAACTCTGGCATGTCATGCCAGTTCCTCTGGACC 3928 101 AGCTGAAGAACTAACTGGAACTGGAGGACTTGGCACTCTGGCACTCTGGACC 160 | 3162 TAACTACAGTACCCATCAAGCAGCATTGAGAAGCTGGGAGTGTACTCTGTGAC 3221 41 TGCTTGTAGGAAAACCTGCAGTCAGACTCCAAACCTCACAGATGCCAG 194 | 3162 TAACTACAGTACCCATCAAGCAGCATTGAGAAGCTGGGAGTGTACTCTGTGAC 3221 41 TGCTTGTAGGAAAACCTGCAGTCAGACTCCAAACCTCACAGATGCCAG 194 |
| Db | 3929 AATACACAGCAGCTGAGGAGGAGACTGACTCTACCTTCTCGAGGACTGTTTACT 4048 161 GTTACATGGAGGAGGACTTGGCACTCTGGCAAGAGCAGCTCCAGCTCTGAAATCTTTA 220 | 3222 AGATGTGGAGGCTCCATTAGTGTCAACCTGTGTCACCTGCCACAGCAGACAT 3281 41 TGCTTGTAGGAAAACCTGCAGTCAGACTCCAAACCTCACAGATGCCAG 254 | 3222 AGATGTGGAGGCTCCATTAGTGTCAACCTGTGTCACCTGCCACAGCAGACAT 3281 41 TGCTTGTAGGAAAACCTGCAGTCAGACTCCAAACCTCACAGATGCCAG 254 |
| Oy | 3989 CCCACTGAGGAGGAGGAGACTGACTCTACCTTCTCGAGGACTGTTTACT 4048 221 TCCACTGAATAGCTGAGGAGCTGACCTTCTCTGAGACTTTCT 280 | 3282 CAAGTGTGG 3292 4049 TCTCAGACGCCGCGCATCCGGCATCTGAGATGCCATGCCCT 400 | 3282 CAAGTGTGG 3292 255 CAAGTGTGG 265 |
| Db | 4169 CCTCTGAGGCCCTAACCTTACCCCTGGGACAGCGGATGCTCCAGCAGGCTG 4228 401 CCTCTGAGGCCCTAACCTTACCCCTGGGACAGCGGATGCTCCAGCAGGCTG 4228 | RESULT 15 AX417823/c LOCUS AX417823 446 bp DNA linear PAT 18-JUN-2002 DEFINITION Sequence 7 from Patent WO231161. ACCESSION AX417823 VERSION AX417823.1 KEYWORDS | RESULT 15 AX417823/c LOCUS AX417823 446 bp DNA linear PAT 18-JUN-2002 DEFINITION Sequence 7 from Patent WO231161. ACCESSION AX417823 VERSION AX417823.1 KEYWORDS |
| Oy | 4229 AGAGAGGCCCGGGCTCTACTGGGCTCTGGCACTGGGAGCTGGCTCTGG 4288 461 AGAACCCCTCCATGACTCTACTGGGAGCTGGCACTGGCTCTGG 520 | REFERENCE AUTHORS Zhu, Z. TITLE Regulation of human phospholipase-like enzyme JOURNAL Patent: WO 02/1161-A 7 18-APR-2002; DBAYER AG (DE) FEATURES source 1..446 Location/Qualifiers /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /db_xref="taxon:9606" | REFERENCE AUTHORS Zhu, Z. TITLE Regulation of human phospholipase-like enzyme JOURNAL Patent: WO 02/1161-A 7 18-APR-2002; DBAYER AG (DE) FEATURES source 1..446 Location/Qualifiers /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /db_xref="taxon:9606" |
| Db | 4289 TGGCATCATCGGACACTGGCTGGAGGTCAGGAGGTGCCGGAGGAGA 4343 521 TGGCATCCTTGGAGACTGTGAGTGGAGAACCGCTCAACAGGAGGA 575 | BASE COUNT 86 a 122 c 147 g 90 t 1 others ORIGIN | BASE COUNT 73 a 80 c 62 g 50 t |
| RESULT 14 | Query Match Best Local Similarity 99.2%; Pred. No. 6.4e-51; Matches 238; Conservative 0; Mismatches 2; Indels 0; Gaps 0; | Query Match Best Local Similarity 99.2%; Pred. No. 6.4e-51; Matches 238; Conservative 0; Mismatches 2; Indels 0; Gaps 0; | Query Match Best Local Similarity 99.2%; Pred. No. 6.4e-51; Matches 238; Conservative 0; Mismatches 2; Indels 0; Gaps 0; |
| AX417822 | AX417822 265 bp DNA linear PAT 18-JUN-2002. | | |
| LOCUS | Sequence 6 from Patent WO231161. | | |
| DEFINITION | | | |
| ACCESSION | | | |
| VERSION | AX417822.1 GI:21522940 | | |
| KEYWORDS | | | |
| SOURCE | Homo sapiens (human) | | |
| ORGANISM | Homo sapiens | | |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| REFERENCE | Zhu, Z. | | |
| AUTHORS | | | |
| TITLE | Regulation of human phospholipase-like enzyme | | |
| JOURNAL | Patent: WO 02/1161-A 7 18-APR-2002; | | |
| DBAYER AG (DE) | | | |
| FEATURES | source | | |
| | 1..446 Location/Qualifiers | | |
| | /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" | | |
| Db | | | |
| FEATURES | | | |
| Source | /organism="Homo sapiens" /mol_type="genomic DNA" | | |

Wed Jan 7 10:03:20 2004

usb-10-054-691-1.rge

Page 22

Search completed: January 6, 2004, 12:19:30
Job time : 15568 secs

APPLICANT: YAMAMOTO, JUN-ICHI
 APPLICANT: ISONO, YUUKO
 APPLICANT: HIO, YURI
 APPLICANT: OTSUKA, KAORU
 APPLICANT: NAGAI, KEIICHI
 APPLICANT: IRIE, RYOTARO
 APPLICANT: TAMECHIKI, ICHIRO
 APPLICANT: SEKI, NAOKO
 APPLICANT: YOSHIKAWA, TSUTOMU
 APPLICANT: OTSUKA, MOTOKU
 APPLICANT: NAGAHARI, KENJI
 APPLICANT: MASUHO, YASUHIKO
 TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
 FILE REFERENCE: 084335/0160
 CURRENT APPLICATION NUMBER: US/10/094,749
 CURRENT FILING DATE: 2002-03-12
 PRIOR APPLICATION NUMBER: 60/350,435
 PRIOR FILING DATE: 2002-01-24
 PRIORITY NUMBER: JP 2001-328381
 NUMBER OF SEQ ID NOS: 381
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 390
 LENGTH: 2477
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-094-749-390

Query Match Best Local Similarity 22.2%; Score 970; DB 13; Length 2477; Matches 973; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1279 AGCTCGCCGGAGATGAGAACATCGGCACCGTACCCCTGGCAACATCCTCGGAA 1338
 Db 6 AGCCGCCGGAGATGAGAACATCGGCACCGTACCCCTGGCAACATCCTCGGAA 65

Qy 1339 TCACACCCCTCCCTGAGGGCTCTCTGTGGCACTGGAAAGAACCGTCTAATGCC 1398
 Db 66 TGCACCCCTCCCTGAGGGCTCTCTGTGGCACTGGAAAGAACCGTCTAATGCC 125

Qy 1399 TCTTTAACAGGGCTGTGCGAGGAGGGCCAGTGAGGAATCTACCTGTGCAAGGCCAGGG 1458
 Db 126 TCTTAACAGGGCTGTGCGAGGAGGGCCAGTGAGGAATCTACCTGTGCAAGGCCAGGG 185

Qy 1459 CTGCTGGACTGTGAGGATGACAGGATACACTTCAGGAGACTGGAGATAATA 1518
 Db 186 CTGCTGGACTGTGAGGATGACAGGAGGATACACTTCAGGAGACTGGAGATAATA 245

Qy 1519 ACCCTGTTTATAGCGGCAATGAGCTCTGTGATTCTGAAATGCTGGTCACTATCT 1578
 Db 246 ACCCTGTTTATAGCGGCAATGAGCTCTGTGATTCTGAAATGCTGGTCACTATCT 305

Qy 1579 CCCAGAAACTTACAGACACATGGAAAGGCCCTGGACATCCCTCATGCTGAGGTCT 1638
 Db 305 CCCAGAACTTACAGACACATGGAAAGGCCCTGGACATCCCTCATGCTGAGGTCT 365

Qy 1639 CGGCATTTGTGACCTGCTGCGTGTGAGATGTCACCTGAGGGAGGTGACAG 1698
 Db 366 CGGCATTTGTGACCTGCTGCGTGTGAGATGTCACCTGAGGGAGGTGACAG 425

Qy 1699 GAGAAAAGTCTACTGCCAGGATGTCCTCGGTCTGTGTCCTCTGTGAG 1758
 Db 426 GAGAAAAGTCTACTGCCAGGATGTCCTCGGTCTGTGTCCTCTGTGAG 485

Qy 1759 TTGATGATACTTACACAGACATGTCACCCATGCAATTCAACAGAGGTTGAG 1818
 Db 486 TTGATGATACTTACACAGACATGTCACCCATGCAATTCAACAGAGGTTGAG 545

Qy 1819 AGACCCACCACTGATGAGGGCATATCACACAGACAGGAGATTTACTGTGTT 1878
 Db 546 AGACCCACCACTGATGAGGGCATATCACACAGACAGGAGATTTACTGTGTT 605

Qy 1879 GTGCAGCCCTTGTGAAACGTGACATGCCAAGACCTCGGAGGATGCTGACAC 1938

RESULT 3

US-09-764-891-1391

Sequence 1391, Application US/09764891

Publication No. US2003007780B1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: P006

CURRENT APPLICATION NUMBER: US/09/764,891

CURRENT FILING DATE: 2001-01-17

PRIOR APPLICATION DATA REMOVED - CONSULT PALM OR FILE WRAPPER

NUMBER OF SEQ ID NOS: 10231

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1391

LENGTH: 572

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (375)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE

LOCATION: (529)

OTHER INFORMATION: n equals a,t,g, or c

US-09-764-891-1391

Query Match Best Local Similarity 9.4%; Score 411.2; DB 11; Length 572; Matches 492; Conservative 2; Mismatches 6; Indels 9; Gaps 7;

Qy 3042 CCTTTGGACCAATATGCTGAAACCTGGAGCAACAGAGACCCCTGGACCTGAGGC 3101
 Db 15 CCTTGAGGACCAATATGCTGAAACCTGGAGCAACAGAGACCCCTGGACCTGAGGC 74

Qy 3102 AGAGATGCGCATCCCTGCTCACTCAGGATGAGCCCTTCCTGAGAACCCCTGGAAATAG 3161
 Db 75 AGAGATGCGCATCCCTGCTCACTCAGGATGAGCCCTTCCTGAGAACCCCTGGAAATAG 134

Qy 3162 TAATCAGCTTCCATCAACGGGAGTGGCCACTGACTCTGTGTCAC 3221
 Db 115 TAATCAGCTTCCATCAACGGGAGTGGCCACTGACTCTGTGTCAC 194

Qy 3222 AGAGTGGAGGCTTCCATAGTGTGTCACCTCTGTGTCACGGGAGGATTTACTGTGTT 3281
 Db 195 AGAGTGGAGGCTTCCATAGTGTGTCACCTCTGTGTCACGGGAGGATTTACTGTGTT 254

QY 3282 CLAAGTGGGGCCCTGGGAACTCTGACTACACCGAGGACTCGACCAACAA 3341
Db 255 CAAAGTGTTGGGCCCTGGGAACTCTGACTACACCGAGGACTCGACCAACAA 314
QY 3342 CTCCAGTC-ACCTTACCCATCTGGGGGACTCTCTT-GGAGCATGGAGGGATGG 3399
Db 315 CTCCAGTC-ACCTTACCCATCTGGGGGACTCTCTT-GGAGCATGGAGGGATGG 374
QY 3400 AA-CTTGAGGTACACCAACTTGGAACATTCTGAGAGGTACACCTTACTCC 3457
Db 375 INACTTGGAGACTCACCAACTGCCAACATTCTGAGAGGTACACCTTACTCC 434
QY 3458 TGGCTTCTT-ACCGAGCTGGG-AGGGGAGGACT-AAATGTCAGCGGAA 3513
Db 435 TGGCTTCTGACAGGACTGGGGAGGGAGCAGCAGGACTGAAATGTCAGCGGAA 494
QY 3514 GGCGCCAGAGTA-GGGCATCAGCCC 341
Db 495 GGCGCCAGAGTAGGGGACATGCCAGCCC 523

RESULT 4
US-10-184-644-290/c
; Sequence 290, Application US/10184644
; Publication No. US2003004930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 290
; LENGTH: 1523
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-184-644-290/c

Query Match 1.0%; Score 42.4; DB 15; Length 1523;
Best Local Similarity 9.8%; Pred. No. 0.066; Matches 83; Conservative 200; Mismatches 567; Indels 0; Gaps 0;

QY 2625 CAACTTGTGACCATCTCCGAAATGCTTGCAGAGGTCCCCAGT 2684
Db 1167 SHMCB.N.BA..M.GY.CT.C..GYGA...BB.G.TCN.G.R.HG.SA...BK...D.G. 1108
QY 2685 CCTGGTCACCTCGTGGACTTCTGAAACCCACTATGCGCAGGTCTGGAAA 2744
Db 1107 .GS.C..GAGAREN.AHKG.TDYGMDTBGHINHA.G.MCSR.C.BG.GS.CNHR..GMT. 1048
QY 2745 CCCAGACAGTCGCCAGTGAGCCGGCGAGGCTTGTGACTGCTTCTGACCTGCG 2804
Db 1047 D.G.N...BGDH.B.HG.CARN.G.GBRNN.CHRNH.GHNH.GY.C. 988
QY 2805 GAGAACTCCAGAGCTAGCCAGGCTTCTGGAGGAGCTTCTGGAAA 2864
Db 987 C..GSGS.CHMHS..DGACCD.G.N..GAN..BAGHMCR.R.GCYR..B.H..AGACN 928
QY 2865 GGCGGAGCTGGGGTCAAGGGCGCTAGACCGAGGACTCTCTGCTGCTGCA 2924

Db 927 NMG.SS.GTINGMB.N.HB.CMG..YDA.AA...YHTK...SSGYT.C..MRCTMBW.S. 868
QY 2925 GCCTCTTCAGAACATCCAGTCCTCCCTGCTCTGGGGATGGCTCCAGATACGTCCTT 2984
Db 867 WY.SGHGD..NAC.T.DS.SA.HN.SC..BSS.HNCD.A.BY.SY.CN.TD..GY.YNR 808
QY 2985 CTTGCCAGACTGCATCCACAAATCAGAAATCCACTCCAGCTGCCAGAGCCT 3044
Db 807 S..AS.DSKNS.ARNA.KS.SNNS.H..A.DY.TS..Y.BTA.DNC..R..ABHM.KCY 748
QY 3045 TGGACCATATGCTGAGACACTTGGAGAGCTGAGACCTGAGAGCAGA 3104
Db 747 ..TY.CMSGYBBA.KDAG..GY.S..GSS..NCHGA.H..TBH....M...M.GY. 688
QY 3105 GATGCCCATCACCTGRCACACTCAGATGACGCCCTCCTGAGAACCCCTCGCAATGATAA 3164
Db 687 NCSB.YMM.WMC.WT.DNGN..NS..N.AS.SB.AA.TC.A.AA.YNHR.S..YBSS. 628
QY 3165 CTACACGTACCCATCAGCCAGCATCTGAGACTGGGCACTGACTCTCTGTGACMA 3224
Db 627 CT.AHNSBGS..NSY.K.AM.CS.CY.BYCDBA..NCA.K..BSPTCH.TC.YB.M.M 568
QY 3225 GTGGAAGGCTCCAATAGTGTCCAACTCTCTTACACAGCTCCGACGAGACATCAA 3284
Db 567 NNS.N.MY.N..MM..CAT..BSB.NHN.Y.HABR...DS..YB.M.NSGIR.AC.GYGM 508
QY 3285 ACTGGTGGCCSCCTGGTGAATCTGACTACAGGTGGAGCTGCCAACACTC 3344
Db 507 ..GB.HK.G.SS.YSYRH.SCSGY.MMSM..S.YMNT.YY.SSGYTCSA..NH..RHT. 448
QY 3345 CACTGACCTACCCACATCTGGGGGACTCTTGAGCACTGGAGGAATGGGAACTT 3404
Db 447 WM.DHGB..N.T.D.A...S...T.CMS.A..MNTR.S..N.N..H..AMBY.GN.MNT 388
QY 3405 GGAGACTCACCAACAGTGCACACATCTGAGAAGTTCACCTTACCTCCTGGCTT 3464
Db 387 N.....SB.CH..CMT..A.MNCR.B..SA.SM.C..TH.T.HS..NMS.H.YM.MR.A 328
QY 3465 CTCTTACAGC 3474
Db 327 .TCT..TM.S 318

RESULT 5
US-10-184-634-290/c
; Sequence 290, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 290
; LENGTH: 1523
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-184-634-290/c

Query Match 1.0%; Score 42.4; DB 15; Length 1523;

Best Local Similarity 9.8%; Pred. No. 0.066; Mismatches 567; Indels 0; Gaps 0; Matches 83; Conservative 200; Mismatches 567; Indels 0; Gaps 0;

QY 2625 CAACTTGTGACCATCTCGGAATGCCATTGAGCTTCATGAGAGCTGGCCAGAT 2684

Db 1167 SHMCB.NBA..M.GY.CT.C..GYA...TCN.G.R.HG.SA..BK..D.G. 1108

QY 2685 CTCTGGTCAACCTCGTGGACTTCTGAGACCCCCATATCATGGCAGGGTTCTGGAAA 2744

Db 1107 .CS.C..GAGARCN.AHBC.TCDYGMDBGHHNHA.G.MCSR.BG.GS.CMH..GMT. 1048

QY 2745 CCCAGACAGTGCCCACTGGCAGAGGGCAGCGRUTTGTAACAGGGTCTGACCCCTGG 2804

Db 1047 D.G.N..BGDH.B.HG..CARN.G.GBRNN.CHBGATNN.GHNH.GHN..GY.C.. 988

QY 2805 GAGAGACTCCCAGAGCTACCCAGGGCTTGGAGGCCTTCAGGGAGCCCTACGGAGCAGAT 2864

Db 987 C..GSGS.CHMDHS..DEACCD.G.N..GAN..BAGHCMR.R.GTGR..B.H.AGACN 928

QY 2865 GCGCGAGCTGGTGGGGTAGGGCCCTATGACACCCAGGAGACTCTCTGGGTGCTCA 2924

Db 927 NMG.SS.GTINGMB.N.HB.CMG..YDA.AA..YHTK..SSGT.C..MRCTMBW.S. 868

QY 2925 GCCTTCTTCAGACATCCAGCTCCAGGAACTCCACTCCAGCTGCCAGAGCCCT 2984

Db 867 WY.SGHGD..NAC.T.DS.SA.HN.SC..BSN.HNCD.A.BY.SY.CN.TDB..GY.YNR. 808

QY 2985 CTTGGCCCCAGACTGCATCCACCCAAATCAGAACTCCACTCCAGCTGCCAGAGCC 3044

Db 807 S..AS.DSKNS.ARNA.XS.SINNS.H..DY.TS..Y.BTA.DNC..R.ABHM.KCY. 748

QY 3045 TTGGACCAAATGCTGTGAAACCCTGGAAGCAGAACAGAGACCTGGACAGAGCA 3104

Db 747 ..TY.CMNSGYBBA.KGAG..GY.S..GSS..NCHGA.H..TBH..M..M.GY. 688

QY 3105 GATGCCCATCACCTGCTCCACTCGAATGAGCCCTTCCTGAGAACCCCTGGATA 3164

Db 687 NCNB.YYMY.WMC.WT.DGENN..NS..N.AS.SB.AA.TC.A.AA.YNHR.S..YBSS. 628

QY 3165 CTACACGACCCATCAAGCCAGCATGGAGACTGGGCAGTGACTTCCTGTCATAGA 3224

Db 627 CT..AHNSBOS..NSY.K.AM..CS.CY.BYCDBA..NCA.K..BSTMCH.TC.YB..M.M 568

QY 3225 GTGGAGGCTTCATAGTGTCAACCTCTGTCACCACTGGAGCTGACATCAA 3284

Db 567 NNS.N.MY.N..MM..CAT..BSB..NHN.Y.HABR..DS..YB.M.NSGHB.AC.GYGM 508

QY 3285 ATGTGTGTGCGCCCTGGGAGCATCTGTCATGAGCTGGAGCTGACCAACACTC 3344

Db 507 ..GB.HK.G.S..YSYRH.SCSGY.MMSM..S.YMNT.YY.SSGYTCSA..NH..RHT. 448

QY 3345 CAGTGACCTACCCACATCTGAGGGACTCTCTGGACATGGAGGGATGGAACTT 3404

Db 447 WM.DGHGB..N.T.D.A..S..T..CMS.A..MNHR.S..N.N..H..ANBY.GN.MNT. 388

QY 3405 GAGAGACTCACACCAACTGCCAACATTGTGAGAAGTCACCTTACTCCITGGCTT 3464

Db 387 N.....SB.CH..CMT..A.MNCR.B..SA..SM.C..TH.T.HS..NMS.H..YM.MNR.A. 328

QY 3465 CTCTTACCAAGC 3474

Db 327 .TCT..TM.S 318

RESULT 6
US-10-031-268

Sequence 268, Application US/10157031
Publication No. US20030108890A1

GENERAL INFORMATION:

APPLICANT: Baranova, A. V.

APPLICANT: Yankovskiy, N. K.

APPLICANT: Kozlov, A. P.

APPLICANT: Lobashev, A. V.

APPLICANT: Krukovskaya, L. L.
TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences

FILE REFERENCE: 2760-103
CURRENT APPLICATION NUMBER: US/10/157,031

CURRENT FILING DATE: 2002-05-30
NUMBER OF SEQ ID NOS: 415

SOFTWARE: PatentIn version 3.1
SEQ ID NO: 268

LENGTH: 1186
TYPE: DNA

ORGANISM: Homo sapiens
US-10-157-031-268

Query Match 1.0%; Score 41.8; DB 13; Length 2400;

Best Local Similarity 61.5%; Pred. No. 0.14; Mismatches 42; Indels 0; Gaps 0; Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 4206 CCCATTGTCCTCCAGACCCGGCTGAGAGGCCCCGAGGTGCTACTGGGCTGTCCTGA 4265

Db 1408 CTGGTACTTCAGACGGGGTTCTCCATGACTCTGGGGCTGTCCTGA 1349

QY 4226 GCGAGCGAGGAGGGCTGAGGGCTGAGAGGCCCCGAGGTGCTACTGGGCTGTCCTGA 4314

Db 1348 CGAGTGGGGCGCTGAGAGGCCCCGAGGTGCTACTGGGCTGTCCTGA 1300

RESULT 8
US-10-172-620-15/C

Sequence 15, Application US/10172620
Publication No. US20030053993A1

GENERAL INFORMATION:

APPLICANT: Hung, Mien-Chie
; APPLICANT: Lin, Shiaw-Yih
; TITLE OF INVENTION: Methods and Compositions for Inhibiting EGR Receptor
; FILE REFERENCE: USNC:20US
; CURRENT APPLICATION NUMBER: US/10/172,620
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/298,579
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 15
; LENGTH: 2643
; TYPE: DNA
; ORGANISM: Human
; US-10-172-620-15

Query Match 1.0%; Score 41.8; DB 15; Length 2643;
Best Local Similarity 61.5%; Pred. No. 0; Mismatches 42; Indels 0; Gaps 0;
Matches 67; Conservative 0;

Qy 4206 CGGATTCCTCCAGACCGAGGTGAAGAGGCCCGAGGTGCCTACTGGCTGTCAGT 4265
Db 1993 CTGGTACTCCAGACCAAGGGTGTGTTCTCCATGACTCTCTGGGGCGGTCTG 1934

Qy 4266 GCGAGCGGAGTCGGCTTGCTGGCATCATCGGACAGCTGCTCG 4314
Db 1933 CCGAGTGGGCCGTCATGATGTCGGCACACTGATGCTCG 1885

RESULT 9

US-09-725-433-1/c

; Sequence 1, Application US/09725433
; Patent No. US20068362A1
; GENERAL INFORMATION:
; APPLICANT: No. US/0020068362A1 artis AG
; TITLE OF INVENTION: Increased transgene expression in retroviral vectors having a sca
; FILE REFERENCE: 4-3092/B/SYS
; CURRENT APPLICATION NUMBER: US/09/725,433
; CURRENT FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3633)
; OTHER INFORMATION:
; US-09-725-433-1

Query Match 1.0%; Score 41.8; DB 9; length 3633;
Best Local Similarity 61.5%; Pred. No. 0; Mismatches 42; Indels 0; Gaps 0;
Matches 67; Conservative 0;

Qy 4206 CGGATTCCTCCAGACCGAGGTGAAGAGGCCCGAGGTGCCTACTGGCTGTCAGT 4265
Db 1834 CTGGTACTCCAGACCAAGGGTGTGTTCTCCATGACTCTCTGGGGCGGTCTG 1775

Qy 4266 GCGAGCGGAGTCGGCTTGCTGGCATCATCGGACAGCTGCTCG 4314
Db 1774 CCGAGTGGGCCGTCATGATGTCGGCACACTGATGCTCG 1726

RESULT 10

US-09-920-300A-1731/c

; Sequence 1, Application US/09920300A
; Patent No. US200203672A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun

Query Match 1.0%; Score 41.8; DB 13; Length 5264;
Best Local Similarity 61.5%; Pred. No. 0; Mismatches 42; Indels 0; Gaps 0;
Matches 67; Conservative 0;

Qy 4206 CGGATTCCTCCAGACCGAGGTGAAGAGGCCCGAGGTGCCTACTGGCTGTCAGT 4265
Db 1753 CTGGTACTCCAGACCAAGGGTGTGTTCTCCATGACTCTCTGGGGCGGTCTG 1694

Qy 4266 GCGAGCGGAGTCGGCTTGCTGGCATCATCGGACAGCTGCTCG 4314
Db 1693 CCGAGTGGGCCGTCATGATGTCGGCACACTGATGCTCG 1645

RESULT 11

US-10-099-926-1731/c

; Sequence 1, Application US/10099926
; Publication No. US2003016606A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Scristi, Heather
; APPLICANT: Jiang, Yudiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121..547C2
; CURRENT APPLICATION NUMBER: US/10/099,926
; CURRENT FILING DATE: 2002-03-17
; NUMBER OF SEQ ID NOS: 1982
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1731
; LENGTH: 5264
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2497
; OTHER INFORMATION: n = A,T,C or G

Query Match 1.0%; Score 41.8; DB 13; Length 5264;
Best Local Similarity 61.5%; Pred. No. 0; Mismatches 42; Indels 0; Gaps 0;
Matches 67; Conservative 0;

Qy 4206 CGGATTCCTCCAGACCGAGGTGAAGAGGCCCGAGGTGCCTACTGGCTGTCAGT 4265
Db 1753 CTGGTACTCCAGACCAAGGGTGTGTTCTCCATGACTCTCTGGGGCGGTCTG 1694

Sequence 1731, Application US10033528
 Publication No. US20020131971A1
 GENERAL INFORMATION:
 APPLICANT: KING, Gordon E.
 APPLICANT: Magher, Madeleine Joy
 APPLICANT: Xu, Jiangchun
 APPLICANT: Sacrist, Heather
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
 FILE REFERENCE: 20121..547C1
 CURRENT APPLICATION NUMBER: US/10/033, 528
 CURRENT FILING DATE: 2001-12-26
 NUMBER OF SEQ ID NOS: 1196
 SOFTWARE: Fast-SEQ for Windows Version 4.0
 SEQ ID NO 1731
 LENGTH: 5264
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 2497
 OTHER INFORMATION: n = A,T,C or G
 US-10-013-528-1731

Query Match 1.0%; Score 41.8; DB 14; Length 5264;
 Best Local Similarity 61.5%; Pred. No. 0.23; Mismatches 0; Indels 0; Gaps 0;
 Matches 67; Conservative

Qy 4206 CGCGATGCTCCAGACAGGGCTGAGAAGCCCCGAGGGTGTACTGGCTGCCAGT 4265
 Db 1753 CTGGTACTCTCAGACAGGGTGTTCTCCATGACTCTGGCTGCCAGT 1694

Qy 4266 GGAGGCGGAGTCGGCTCTGGCTCATCGGACAGGGCTCG 4314
 Db 1693 CGGAGTGGGGCCGTCATGTAGGGCACATCGGATACAGTTCCTGG 1645

RESULT 13
 US-10-007-926A-137/c
 Sequence 137, Application US/10007926A
 Publication No. US20030143539A1
 GENERAL INFORMATION:
 APPLICANT: BERTUCCI, FRANCOIS
 APPLICANT: HODIGATTE, REMI
 APPLICANT: BIRNBAUM, DANIEL
 APPLICANT: NGUYEN, CATHERINE
 APPLICANT: VIENS, PATRICE
 APPLICANT: FEERT, VINCENT
 TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
 TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
 FILE REFERENCE: 1556-R-00
 CURRENT APPLICATION NUMBER: US/10/007, 926A
 CURRENT FILING DATE: 2001-12-07
 PRIOR APPLICATION NUMBER: 60/254, 090
 PRIOR FILING DATE: 2000-12-08
 NUMBER OF SEQ ID NOS: 468
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 137
 LENGTH: 5532
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: epidermal growth factor receptor (avian v-erb-b) oncogene
 OTHER INFORMATION: erythroblastic leukemia viral (v-erb-b) oncogene
 OTHER INFORMATION: homolog (EGFR) gene.
 US-10-007-926A-137

Query Match 1.0%; Score 41.8; DB 13; Length 5532;
 Best Local Similarity 61.5%; Pred. No. 0.23; Mismatches 0; Indels 0; Gaps 0;
 Matches 67; Conservative

Qy 4206 CGGATGCTCCAGACAGGGCTGAGAAGCCCCGAGGGTGTACTGGCTGCCAGT 4265
 Db 2020 CGGGTACTCTCCAGACGGGGTGTCTCCATGACTCCIGCGGGAGGTCTGA 1961

RESULT 14
 US-10-101-510-95/c
 Sequence 95, Application US/10101510
 Publication No. US20030148295A1
 GENERAL INFORMATION:
 APPLICANT: WAN, JACKSON
 APPLICANT: WANG, YIXIN
 TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
 FILE REFERENCE: 1117.0012
 CURRENT APPLICATION NUMBER: US/10/101, 510
 CURRENT FILING DATE: 2002-03-20
 PRIORITY NUMBER: 60/276, 947
 PRIORITY FILING DATE: 2001-03-20
 NUMBER OF SEQ ID NOS: 805
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 95
 LENGTH: 5532
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-101-510-95

Query Match 1.0%; Score 41.8; DB 13; Length 5532;
 Best Local Similarity 61.5%; Pred. No. 0.23; Mismatches 0; Indels 0; Gaps 0;
 Matches 67; Conservative

Qy 4206 CGGATGCTCCAGACAGGGCTGAGAAGCCCCGAGGGTGTACTGGCTGCCAGT 4265
 Db 2020 CGGGTACTCTCCAGACGGGGTGTCTCCATGACTCCIGCGGGAGGTCTGA 1961

RESULT 15
 US-10-380-931-17/c
 Sequence 17, Application US/10380931
 Publication No. US20030215944A1
 GENERAL INFORMATION:
 APPLICANT: ISIB Pharmaceuticals, Inc.
 APPLICANT: C. Frank Bennett
 APPLICANT: Jacqueline Wyatt
 APPLICANT: Susan M. Freier
 TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HER-1 EXPRESSION
 FILE REFERENCE: RUSP-0187
 CURRENT APPLICATION NUMBER: US/10/380, 931
 CURRENT FILING DATE: 2003-03-18
 PRIOR APPLICATION NUMBER: 09/676, 610
 PRIOR FILING DATE: 2000-09-29
 NUMBER OF SEQ ID NOS: 182
 SEQ ID NO 17
 LENGTH: 5532
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (187)...(3819)
 US-10-380-931-17

Query Match 1.0%; Score 41.8; DB 13; Length 5532;
 Best Local Similarity 61.5%; Pred. No. 0.23; Mismatches 0; Indels 0; Gaps 0;
 Matches 67; Conservative

Qy 4206 CGGATGCTCCAGACAGGGCTGAGAAGCCCCGAGGGTGTACTGGCTGCCAGT 4265
 Db 2020 CGGGTACTCTCCAGACGGGGTGTCTCCATGACTCCIGCGGGAGGTCTGA 1961

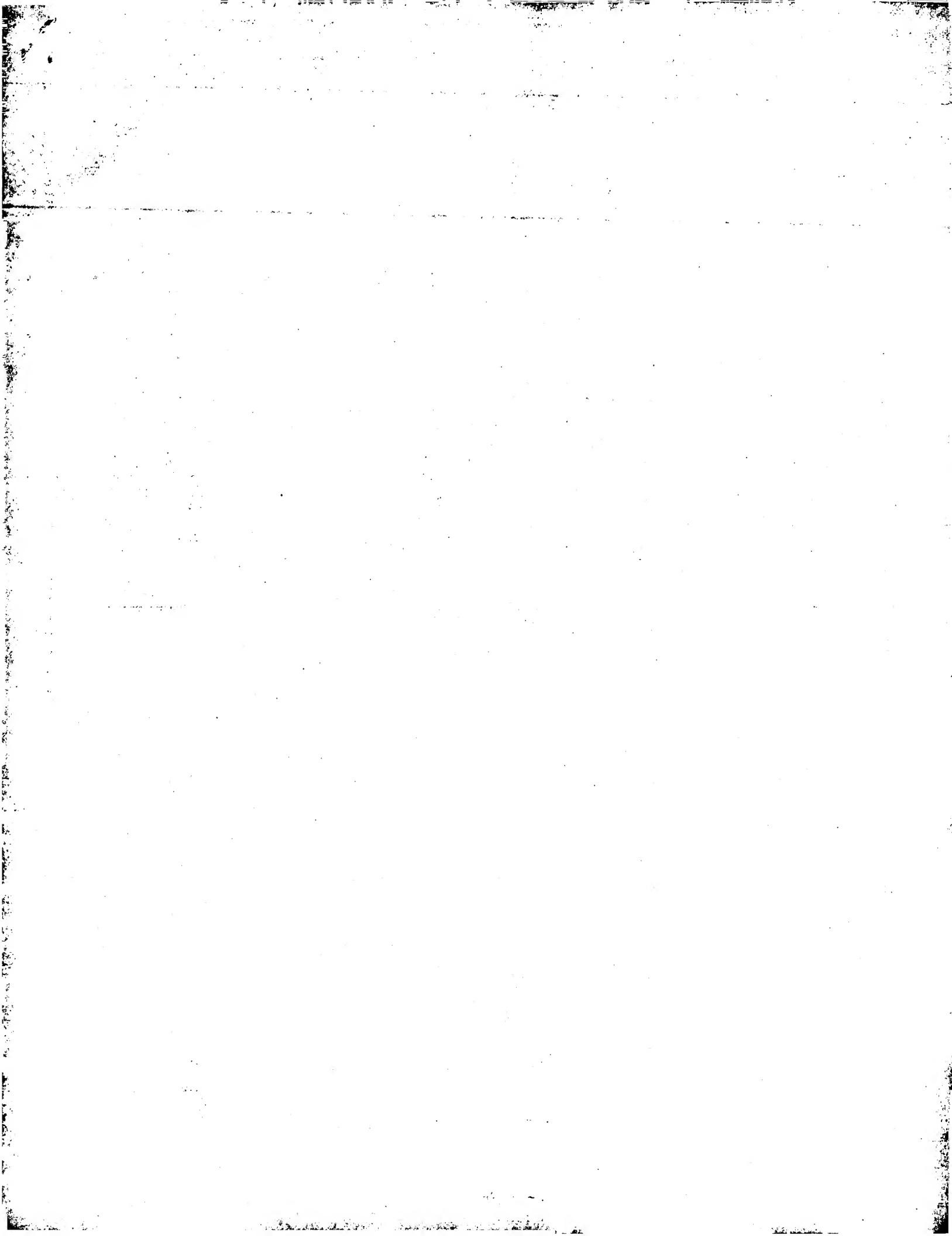
Wed Jan 7 10:03:21 2004

us-10-054-691-1.rnpb

Page 9

Qy 4266 GGAGCGGAGTCGGCTTGTTGGCATCATCGGACAGTGTCTGG 4314
Db 1960 CGCAGTGGCGGTCAATGTACTGGCACATGATGTCGG 1912

Search completed: January 6, 2004, 18:54:58
Job time : 1316 secs



QY 4206 CGATTCGCTCCAGACCGAGCTGAAGAAGCCCCAGGTCTTACTGGCTGTCCTCACTG 4265
 Db 2020 CTGGTACTCTCACACCAAGGTGTTGTTCTCCATGACTCTTGCCGGAGTCCTGA 1961
 QY 4266 GGCAGCGGGAGTCGCCTGTGGTAGGCATATCGGACAGTCCTGG 4314
 Db 1960 CGACTGGGGCCGTCATACTAGTGGCACACTGATACTGTTCTGG 1912

RESULT 2
 US-09-676-610B-17/c
 ; Sequence 17 : Application US/09676610B
 Patent No. 644465
 GENERAL INFORMATION:
 APPLICANT: C. Frank Bennett
 APPLICANT: Jacqueline Watt
 APPLICANT: Susan M. Fraier
 TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HER-1 EXPRESSION
 FILE REFERENCE: RTS-0138
 CURRENT APPLICATION NUMBER: US/09/676,610B
 CURRENT FILING DATE: 2000-09-29
 NUMBER OF SEQ ID NOS: 182
 SEQ ID NO 17
 LENGTH: 5532
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (187) ... (3819)

US-09-676-610B-17

Query Match 1.0%; Score 41.8; DB 4; Length 5532;
 Best Local Similarity 61.5%; Pred. No. 0.27%;
 Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 4206 CGATTGCTCCAGACCGAGCTGAAGAAGCCCCAGGTCTTACTGGCTGTCCTCACTG 4265
 Db 2020 CTGGTACTCTCACACCAAGGTGTTGTTCTCCATGACTCTTGCCGGAGTCCTGA 1961
 QY 4266 GGCAGCGGGAGTCGCCTGTGGTAGGCATATCGGACAGTCCTGG 4314
 Db 1960 CGCAGTGCGGCCCTCAATGTTGGGACATCGATACAGTGTCTGG 1912

RESULT 3
 US-09-676-610B-24/c
 ; Sequence 24 : Application US/09676610B
 ; General Information:
 APPLICANT: C. Frank Bennett
 APPLICANT: Jacqueline Watt
 APPLICANT: Susan M. Fraier
 TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HER-1 EXPRESSION
 FILE REFERENCE: RTS-0138
 CURRENT APPLICATION NUMBER: US/09/676,610B
 CURRENT FILING DATE: 2000-09-29
 NUMBER OF SEQ ID NOS: 182
 SEQ ID NO 24
 LENGTH: 169998
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: exon
 LOCATION: (1208) ... (1472)
 NAME/KEY: intron
 LOCATION: (1473) ... (124390)
 NAME/KEY: exon
 LOCATION: (124391) ... (124544)
 LOCATION: (12545) ... (125409)
 NAME/KEY: exon
 LOCATION: (125410) ... (125595)
 NAME/KEY: intron

US-09-676-610B-24

Query Match 1.0%; Score 41.8; DB 4; Length 169998;
 Best Local Similarity 61.5%; Pred. No. 2.1%;
 Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

LOCATION: (125596) ... (12871)
 NAME/KEY: exon
 LOCATION: (128712) ... (128848)
 NAME/KEY: intron
 LOCATION: (128849) ... (133400)
 LOCATION: (133401) ... (133469)
 LOCATION: (134774) ... (136116)
 NAME/KEY: intron
 LOCATION: (133470) ... (134652)
 NAME/KEY: exon
 LOCATION: (134653) ... (134773)
 LOCATION: (136262) ... (137936)
 NAME/KEY: exon
 LOCATION: (137937) ... (138053)
 NAME/KEY: exon
 LOCATION: (136117) ... (136263)
 NAME/KEY: intron
 LOCATION: (138054) ... (138637)
 NAME/KEY: exon
 LOCATION: (138638) ... (138766)
 NAME/KEY: exon
 LOCATION: (138767) ... (138864)
 NAME/KEY: exon
 LOCATION: (138865) ... (138940)
 NAME/KEY: intron
 LOCATION: (138941) ... (139765)
 NAME/KEY: intron
 LOCATION: (139766) ... (139860)
 NAME/KEY: exon
 LOCATION: (139861) ... (142245)
 NAME/KEY: exon
 LOCATION: (142246) ... (142445)
 NAME/KEY: intron
 LOCATION: (142446) ... (143605)
 NAME/KEY: exon
 LOCATION: (139861) ... (142245)
 NAME/KEY: exon
 LOCATION: (143606) ... (143738)
 NAME/KEY: intron
 LOCATION: (143739) ... (145838)
 NAME/KEY: exon
 LOCATION: (145839) ... (145931)
 NAME/KEY: intron
 LOCATION: (145932) ... (147385)
 NAME/KEY: exon
 LOCATION: (147386) ... (147544)
 NAME/KEY: intron
 LOCATION: (147545) ... (153274)
 NAME/KEY: exon
 LOCATION: (153275) ... (153321)
 NAME/KEY: intron
 LOCATION: (153322) ... (155088)
 NAME/KEY: exon
 LOCATION: (155089) ... (155231)
 NAME/KEY: exon
 LOCATION: (155232) ... (156025)
 NAME/KEY: exon
 LOCATION: (156026) ... (156151)
 NAME/KEY: intron
 LOCATION: (156152) ... (156826)
 NAME/KEY: exon
 LOCATION: (156827) ... (156928)
 NAME/KEY: intron
 LOCATION: (156929) ... (163399)
 NAME/KEY: exon
 LOCATION: (163400) ... (163586)

Qy 4286 CCCATTGTCAGACCCAGGGCTGAGAAGGCCCGAGGTCTACTGGCTTCCCAGT 4265
Db 147498 CTGGTACTTCAGACCCAGGGCTGTTCTCCATGACTCTGCCCCAGGTCTGTA 147439
Qy 4286 GCGAGGGAGTCGGCTTGTTGGGATCATGGACAGCTGCTCG 4314
Db 147438 CCCAGTGGGGCTCAATGTTGTTGGGACACATGGATACAGTGTCTCG 147390
RESULT 4
; Patent No.: US-09-877-177A-10/C
; Sequence: 652919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberge
; TITLE OF INVENTION: Method of determining Epidermal Growth Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: Factor Receptor and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 197496
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-877-177A-10

Query Match 1.0%; Score 41.8; DB 4; Length 197496;
Best Local Similarity 61.5%; Pred. No. 2-3; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 42; Gaps 0;

Qy 4206 CGGATTCCTCCAGACCAAGGCTGAAAGAAGCCCGAGGTCTACTGGCTTCCAGT 4265
Db 155498 CTCGGTACCTCCAGACCAAGGCTGAAAGAAGCCCGAGGTCTACTGGCTTCCAGT 155439
Qy 4266 GGGAGGGAGTGGCTGTTGGGATCATGGGACAGGGTCTCG 4314
Db 155438 CGCAGTGGGGCTCAATGTTGGGACACATGGATACAGTGTCTCG 155390
RESULT 5
; Sequence 34, Application US/0910533A
; Patent No. 6285202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600-438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 36778
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
; US-09-105-537-5

Query Match 0.9%; Score 40.8; DB 3; Length 36778;
Best Local Similarity 45.8%; Pred. No. 1-6; Indels 0; Gaps 0;
Matches 141; Conservative 0; Mismatches 167; Gaps 0;

Qy 3986 TCACCCACTGACGAGAGGGACCTGACCTCTCTCGAGGACTGTTTC 4045
Db 29578 TCACTCTACTCCCGAACCTGGCCACGGCCATCTCC 29637
Qy 4046 ACTCTGAGACGGGGCATGGAGATGGCATCGACTCTGGACACATGCTGAC 4105
Db 29838 CCACCGACCGCCACCCCGACSTCCACCTPACGCCCTCCAGGCCTCT 29697
Qy 4106 CAGTGGCCGAAGACTACCTCCAACACTCTACCCACAGCGAGCAAACCTCAAGTGC 4165
Db 29898 GACTGCAAGCTCGCCCCACAGCCCGACAGCTGGTTACGCCCTGACTGGA 29757
Qy 4166 CCTCTCTGAGAGCCCTACCTCACCCCTGGAGACGGATGTTCCAGACAG 4225
Db 29758 AGCGCTGACGGCCTCGGCCAGGGACCTCTCCGGGATGCTGCGCTCGGA 29817
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
; US-09-105-537-34

Query Match 0.9%; Score 40.8; DB 3; Length 4689;
Best Local Similarity 45.8%; Pred. No. 0-47; Mismatches 167; Indels 0; Gaps 0;
Matches 141; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

Qy 3986 TCACTCCACTGAGAGAGGGACCTGACCTCTCTCCAGGACTGTTTC 4045
Db 2988 TCACTCCACTGCGGAGCTCTGGCCACAGGGCTACCATCTGCTGGCCATCTCC 2647
Qy 4046 ACTCTGAGACGGGGCATGGAGATGGCTACAGTGTCTGGCTGAGCTGAC 4105
RESULT 6
; Sequence 5, Application US/0910553A
; Patent No. 6285202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600-438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 36778
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
; US-09-105-537-5

Query Match 0.9%; Score 40.8; DB 3; Length 36778;
Best Local Similarity 45.8%; Pred. No. 1-6; Indels 0; Gaps 0;
Matches 141; Conservative 0; Mismatches 167; Gaps 0;

Qy 3986 TCACCCACTGACGAGAGGGACCTGACCTCTCTCGAGGACTGTTTC 4045
Db 29578 TCACTCTACTCCCGAACCTGGCCACGGCCATCTCC 29637
Qy 4046 ACTCTGAGACGGGGCATGGAGATGGCATCGACTCTGGACACATGCTGAC 4105
Db 29838 CCACCGACCGCCACCCCGACSTCCACCTPACGCCCTCCAGGCCTCT 29697
Qy 4106 CAGTGGCCGAAGACTACCTCCAACACTCTACCCACAGCGAGCAAACCTCAAGTGC 4165
Db 29898 GACTGCAAGCTCGCCCCACAGCCCGACAGCTGGTTACGCCCTGACTGGA 29757
Qy 4166 CCTCTCTGAGAGCCCTACCTCACCCCTGGAGACGGATGTTCCAGACAG 4225
Db 29758 AGCGCTGACGGCCTCGGCCAGGGACCTCTCCGGGATGCTGCGCTCGGA 29817
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
; US-09-105-537-34

Query Match 0.9%; Score 40.8; DB 3; Length 4689;
Best Local Similarity 45.8%; Pred. No. 0-47; Mismatches 167; Indels 0; Gaps 0;
Matches 141; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

Qy 3986 TCACTCCACTGAGAGAGGGACCTGACCTCTCTCCAGGACTGTTTC 4045
Db 2988 TCACTCCACTGCGGAGCTCTGGCCACAGGGCTACCATCTGCTGGCCATCTCC 2647
Qy 4046 ACTCTGAGACGGGGCATGGAGATGGCTACAGTGTCTGGCTGAGCTGAC 4105
RESULT 7
; Sequence 19, Application US/09320878A
; Patent No. 6117659

; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: McDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A.
; CURRENT FILING DATE: 1998-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1998-05-28
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1998-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 19
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
; US-09-320-878-19
; Query Match 0.9%; Score 40.8; DB 3; Length 38506;
; Best Local Similarity 45.8%; Pred. No. 1.7; Mismatches 0; Indels 0; Gaps 0;
; Matches 141; Conservative 45.8%; Pred. No. 1.7; Mismatches 0; Indels 0; Gaps 0;
; Matches 141; Conservative 45.8%; Pred. No. 1.7; Mismatches 0; Indels 0; Gaps 0;
; Qy 3986 TCACCCACTGAACTGGAGGGACACTGACCTCACCTTCTCGAGGATGTTTC 4045
; Db 27720 TCACCTCAGCAGCGGGATGGCGAGATGGCATCGACTCTGGAGCATCTCC 27779
; Qy 4046 ACTTCTCAGACCGGGCATGCCAGATGGCGCATCGACTCTGGACACATCTGGAC 4105
; Db 27780 CCACCGCAACGGCCACACCCGAGCTCCACCTACGCCCTCCAGACGGGCTCT 27839
; Qy 4106 CACTGGGCCGAGACTTCTCACCACTAACCTCACCCACAGCAGGCCAAACTCAGTGC 4165
; Db 27840 GGCTGAGAGCTCGCGGCCACCGCGCGGAGCTGGCTAACCTGGAG 27899
; Qy 4156 CCTCTCTGAGGCCCTTACCTCTACCCCTGGAGCAGCGATCTCCAGACAG 4225
; Db 27900 AGCGCTTGTACGGCCCTCCGGCCACCGCGGAGCTGGCTGGAG 27959
; Qy 4226 CTGAGAGGCCCTGGAGGTCTACTGGCTGTCCACTGGCGAGTCGGCTTG 4285
; Db 27950 GCGAGCCAGAACGGGAGCTGGCTGGCGGAGGGTGGAC 28019
; Qy 4286 TGGGGGC 4293
; Db 28020 TGGAGCC 28027
; RESULT 8
; US-09-141-908-1
; Sequence 1, Application US/09141908
; Patent No. 650341
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: McDANIEL, Robert
; APPLICANT: TANG, Li
; APPLICANT: BETLACH, Mary
; APPLICANT: McDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
; Title of Invention: Modular PKS Gene Cluster as Scaffold
; FILE REFERENCE: 300622002100
; CURRENT APPLICATION NUMBER: US/09/657,440
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320,878
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 19
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
; US-09-657-440-19
; Query Match 0.9%; Score 40.8; DB 4; Length 38506;

RESULT 15
 US-09-252-991A-5839/c
 Sequence 5839, Application US/09252991A
 ; Patent No. 6511795
 ; GENERAL INFORMATION:
 ; APPLICANT: Mat J. Rubenstein et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107195_136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074, 788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094, 190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO: 5839
 ; LENGTH: 1077
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-5839

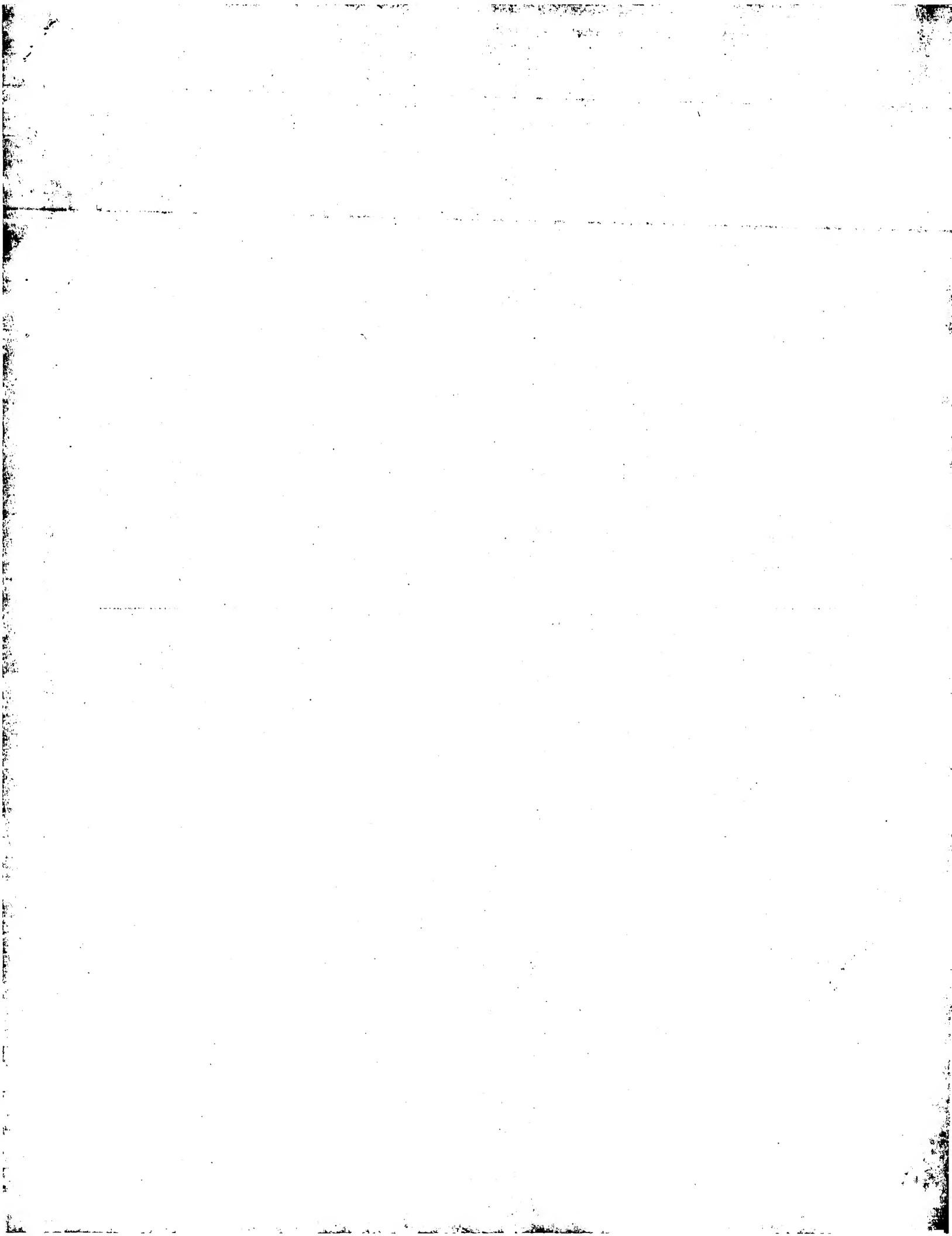
Query Match 0.9%; Score 37.6; DB 4; Length 1077;
 Best Local Similarity 49.5%; Pred. No. 1.6; Mismatches 99; Indels 0; Gaps 0;
 Matches 97; Conservative 0;

```

Qy      2850 CTACCGGAGCACATGCCGAGACTGGTGGGGCAGGCCTATGACACGGAGGACTT 2909
Db      184 CGTGCCGAACCGCTGGTATCGGATGCCAACGCCATACGGAAATCTTCGGCACGT 243
Qy      2910 CTCCTGCTGCTGCAAGCTCTCTCCAGAACATCCAGCTCCCTGGCTGGGGCT 2969
Db      244 GCGGCTGCTGGCGAGCTTCATCTGTACTCTGTTACCTGCTGCTGGGGCT 2969
Qy      2970 CCCAGATACTGCTTC 2985
Db      304 CCTGAGACCTGGTTC 319

```

Search completed: January 6, 2004, 14:45:59
 Job time : 243 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 6, 2004, 06:31:30 ; Search time 8433 Seconds
Scoring table: IDENTITY_NUC

Perfect score: US-10-054-691-1
Sequence: 1 atggggctggggcaggat.....tgccgactgtggccctctag 4377
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0 Maximum DB seq length: 200000000
Post-processing: Minimum Match 0% Maximum Match 100% Listing First 45 summaries

Database : EST: *
1: em_esba: *
2: em_esbum: *
3: em_esin: *
4: em_esmu: *
5: em_esov: *
6: em_espl: *
7: em_esro: *
8: em_htc: *
9: gb_es1: *
10: gb_et2: *
11: gb_hcc: *
12: gb_elt3: *
13: gb_elt4: *
14: gb_elt5: *
15: em_eltfun: *
16: em_eltom: *
17: em_gbs_hum: *
18: em_gbs_inv: *
19: em_gbs_pln: *
20: em_gbs_vrt: *
21: em_gbs_fun: *
22: em_gbs_mam: *
23: em_gbs_mus: *
24: em_ges_pro: *
25: em_ges_rid: *
26: em_ges_png: *
27: em_ges_vri: *
28: gb_gst: *
29: gb_gst2: *

Score: 4377 ; (without alignments)
Searched: 1215238056 residues
Title: US-10-054-691-1
Sequence: 1 atggggctggggcaggat.....tgccgactgtggccctctag 4377
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0 Maximum DB seq length: 200000000
Post-processing: Minimum Match 0% Maximum Match 100% Listing First 45 summaries

Database : EST: *
1: em_esba: *
2: em_esbum: *
3: em_esin: *
4: em_esmu: *
5: em_esov: *
6: em_espl: *
7: em_esro: *
8: em_htc: *
9: gb_es1: *
10: gb_et2: *
11: gb_hcc: *
12: gb_elt3: *
13: gb_elt4: *
14: gb_elt5: *
15: em_eltfun: *
16: em_eltom: *
17: em_gbs_hum: *
18: em_gbs_inv: *
19: em_gbs_pln: *
20: em_gbs_vrt: *
21: em_gbs_fun: *
22: em_gbs_mam: *
23: em_gbs_mus: *
24: em_ges_pro: *
25: em_ges_rid: *
26: em_ges_png: *
27: em_ges_vri: *
28: gb_gst: *
29: gb_gst2: *

ALIGNMENTS

RESULT 1
AK030142
LOCUS AK030142 3148 bp mRNA linear HTC 05-DEC-2002
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493270318 product:weakly similar to PHOSPHOLIPASE B (Rattus norvegicus), full insert sequence.
ACCESSION AK030142
VERSION 1
KEYWORDS HTC; CAP trapper.
SOURCE
ORGANISM Mus musculus (house mouse)
REFERENCE
AUTHORS Bukeryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 9327923
PUBMED 10349636

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|-------|--------------------|-------------------|--------------------|
| 1 | 1626 | 37.1 | 3148 11 AK030142 | AK030142 Mus muscu |
| 2 | 868.8 | 19.8 | 9 935 9 AL542856 | AL542856 AL542856 |
| 3 | 78.8 | 18.0 | 10 12 BM92066 | BM92066 AGENCOURT |
| 4 | 766.4 | 17.5 | 9 935 13 BU150503 | BU150503 AGENCOURT |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| | Source |
|-----------|---|
| REFERENCE | 1. 3148 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref=PRIMERDB:4932703L18" /clone="4932703L18" /sex="male" /tissue_type="testis" /clone_liber=Riken Full-length enriched mouse cDNA library" /dev_stage="adult" |
| AUTHORS | Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamamoto,I., Saito,T., Okazaki,Y., Gotohori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batzloff,S., Casavant,T., Fleischmann,W., Gaasterland,T., Giesi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuoka,Y., Nikaido,I., Peabole,G., Suzuki,R., Tomita,M., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Wagner,L., Washio,T., Sakai,K., Okada,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletchler,C., Fujita,M., Garibaldi,M., Gustincich,S., Hill,D., Hoffmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seva,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hassegawa,Y., Kawaji,H., Kohtsuki,S., and Hayashizaki,Y. |
| TITLE | Functional annotation of a full-length mouse cDNA collection |
| JOURNAL | Nature 409 (6821), 685-690 (2001) |
| MEDLINE | 21053913 |
| PUBLMED | 11076861 |
| REFERENCE | 5 |
| AUTHORS | The FANTOM Consortium and the RIKEN Genome Exploration Research |
| TITLE | Group Phase I & II Team. |
| JOURNAL | Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs |
| MEDLINE | 21053913 |
| PUBLMED | 11217851 |
| REFERENCE | 6 (bases 1 to 3148) |
| AUTHORS | Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayasida,K., Hayatsu,N., Hiramoto,K., Hirayama,T., Hirozane,T., Horie,F., Imotani,K., Ishii,Y., Itoh,M., Murata,M., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsumura,T., Miyazaki,A., Murata,M., Nakamura,M., Nishii,K., Nomura,K., Numazaki,R., Ohno,M., Ohshiro,N., Okada,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sanoo,H., Sasaki,D., Shibata,K., Shirakawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tanaka,A., Takahashi,F., Takuwa,Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M., and Hayashizaki,Y. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0055, Japan (E-mail: genome.res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: +81-45-503-9216, Fax: +81-45-503-9216) |
| COMMENT | CDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken. Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/. |
| FEATURES | Location/Qualifiers |

| | | | | | | | |
|----|------|---|----------|--|--------------|---|-------------------------|
| Qy | 817 | CGGGAGTCCTAACCGTGGTTCCAGCTTCTATGACCAACCCATCTCAC | 876 | Oy | 1897 | AACGTGGACATCCAAGAACCTGGAAAGATGCCCAGAACACTTCTCCCTCTGAC | 1956 |
| Db | 859 | CATGAGACCTCTGGGGTGTTCAGCCTTCTCGATGAATTGACC-----ACCC | 912 | Db | 1933 | AATGTGTCACGCCAGGACCCGGGGCTGGCGACAGTCCTTCTGGGACAATATG | 1992 |
| Qy | 877 | TGGAGGACCCCGACTCCAGATTCTTACCAAGCTGCTGGATAGATG | 936 | Oy | 1957 | TGTTTCACTTCAGCAGCAAAGCTCACTCCGAGCAGCCAGTGCTCTTGGAACATAATG | 2016 |
| Db | 913 | TGAAAAGTCTCGCCCAAGATCCACACACTCGCCCTAGGATCTGGATAGATG | 972 | Db | 1993 | TGTTTCACTTCAGCAGCAAAGCTCACTCCGAGCAGCCAGTGCTCTTGGAACATAATG | 2052 |
| Qy | 937 | ATGGAGCCGGAGAAAGATGAGCATCTAGTGTAAACACGGGGCCATGAG | 996 | Oy | 2017 | CTGGASCTGTTGGCGAGAGACCTGCTAAGTTGAAACAGATGATATACA | 2076 |
| Db | 973 | ATGAAACCTGAGTGTCAAAAGATGSGCTACTCAATCTGAGAAGAAACATGAA | 1032 | Db | 2053 | CTGAAACCTGTTGGCCAAAGACAGAAATAATTTGAATCAGGCCATATGTG | 2112 |
| Qy | 997 | TGTCCTCTCAGAGAGGCCTATCTGTCACTACAGAAACAGCACTACGTGACCA | 1056 | Oy | 2077 | TGTCGACCCAGGTCCAGGGTTCTGGAGCTACAGAACAGCATGCAAGGTATGG | 2136 |
| Db | 1033 | TGTCCTCTGAGAGATGCGCTTCTCTCTACTACAAATAAGCACTACAGGCGA | 1092 | Db | 2113 | TGTCGACCCAGGCTTCAGGACATGAGAGAGGGCTGACCTACAGCAACTGGTAATG | 2172 |
| Qy | 1057 | CTGAGAACCCCAAGAACAGCTGAGCTTAACTGAGAAGAGGGAAATCAGATCTGAC | 1116 | Oy | 2137 | ACCTGGTGCATGAGACAGAGAGGGCTGACCTCTGCCTTGACCCCTACACTGATGCC | 2196 |
| Db | 1093 | CTGAGAACCTCTGATAGGTTCACCTCACTGAGTGAAGAAGAGGGATGAGATG | 1152 | Db | 2173 | ACCTGGATGGTGCAGAGAGAGGGCTGACCTCTGCCTTGACCCCTACACTGATGCC | 2232 |
| Qy | 1117 | ATAGACCCCTGAGTAACTACATCCCAACAGTTCAGTGAAGAAGAGGGATGAGATG | 1176 | Oy | 2197 | CTGAGACCTGTTGGCGAGCTTCAAGTGTGAGCTGAGTGTGAGATGAGATG | 2256 |
| Db | 1153 | AGAAACCCCTCAACTACATCCCAACAGTTCAGTGAAGTACGCTTCACTGAG | 1212 | Db | 2233 | CTGAGACCTGAGACATCTAAGTTGTCAGTGTGAGCTTCACTGAGT | 2292 |
| Qy | 1177 | GTAATTGGAGCCCTGGGACTCTCAAGGGAGGAATGGGCTCAACCTGG | 1236 | Oy | 2257 | GNATTCGCTCAACAGAACCCATCCCGATGTCACACAGCATGTCAGGGTCA | 2316 |
| Db | 1213 | ATCATAGAGGAGCTGGGTGACCGCTCACGGAGGAATGGACTGGGGCGCCCTGG | 1272 | Db | 2293 | GNATTCGCTCCAGGCTTCGACCATTTCTGAGCACTACAGCAACTGGTAATG | 2352 |
| Qy | 1237 | ATACATGGTCTGAGCTCTGACTCTAGTACCGAGGCTGCTGAGATG | 1296 | Oy | 2317 | TACAGTGAGGGAGGGCTCCTGGAGATGTGACCACTTACCTATATCCTGAG | 2376 |
| Db | 1273 | ATATATCTTGGATGCTTAACTGAAATACCGAGGCTGCTGAGTGTGGGGATG | 1332 | Db | 2353 | TATAGTGCTGGTGGAGAACGACCTCCGGATGTCACACTTGGCAATTCAGG | 2412 |
| Qy | 1297 | ACATCGGACCGTTACACCTGGGACATCTCGGAAATTCAACCTTCCCTGAG | 1356 | Oy | 2377 | GAGTTTACAGAAACCTCACAGGCTACGGCGGCGGGACGGTGTGACGAT | 2436 |
| Db | 1333 | ACATCAAAACCGTGCACCCATTACCCACATCTCCGGGAATTCACCCCTCTGAG | 1392 | Db | 2413 | AAATTATGAAATCTACAGGCTACTCAGTAGGAAACCGGTACTCCGGCAAC | 2472 |
| Qy | 1357 | GGCTCTCTGTTGGACTGGAAAGAACAGCTCTTAATGCTTCTTAACAGGTGTG | 1416 | Oy | 2437 | GCATTCCTCAATCAAGTGTCCGGAGAACGAAAGCTCTGGAGCTTGTGACCA | 2496 |
| Db | 1393 | GCCTTCTCTGTTGGACTGGAAAGAACGACTTCCGGACATCTCAGGCGTA | 1452 | Db | 2473 | GCTTCTTAACCAAGGCTGTCCTGGGGGAGGCTGAGACCTTACGGCAAGTCCC | 2532 |
| Qy | 1417 | GAAGGAGCCGAGCTGAGGATCTACCTGGCCAGGAGCTGGTGGACCTGATGAG | 1476 | Oy | 2497 | ACTCTGAGCAGAGATGAAAGATGATGATGAGTAAATTCATGAGAACTGGAAGTC | 2556 |
| Db | 1453 | GAAGGAGCCAAATCTGATGGCTTAGCTGGCAAGGCCAGAAACTGGGAG | 1512 | Db | 2533 | ACTCTGATGAGATGAGACTGAGTGTGAGCTTCACTGAGCTTGTGAGTC | 2592 |
| Qy | 1477 | ATGACACAGAGGATACCTTCTGGAGAGCTGGAGATAATACCTGTTAGGGC | 1536 | Oy | 2557 | ATCAGCTGCTGATCTGGGAGGTTCTGGGAGGTTCTGGGAGGTTCTGGGAGG | 2616 |
| Db | 1513 | GAAGACAGACATAAACTTCTGGAGAGCTGGAGATAATACCTGTTAGGGC | 1572 | Db | 2593 | ATCAGGTTGATGATCTGGGGAGGTTCTGGGAGGTTCTGGGAGGTTCTGGGAGG | 2652 |
| Qy | 1537 | ATGACCTCTGTTGATTCTGCAATGATCTGGCACTATTCTCCCCAGAACTCACAGAC | 1596 | Oy | 2617 | TCTGACCCAACTTGTGAGCATCTGGCAATGCTTGGAGCTCTGGATAGAGGTG | 2676 |
| Db | 1573 | ATGACCTCTGTTCTGGCTTCTGGAGTAATCTGAGCTGCGCTTCTCCCAAACCTCATGAG | 1632 | Db | 2653 | TGTGAGCAATTCTGGTACATCTCCAAATGCTGGATCTCATAGAGGAGTA | 2712 |
| Qy | 1597 | ACATGGAGACCCCTGGAGACTCTCCATGAGCTGGAGTTCTGGGAGATTGTGAGCTG | 1656 | Oy | 2677 | CCGAGCTCTGGCAACCTCTGGAGCTTCTGGAGACCCACATCTAGCGCAGGTTTC | 2736 |
| Db | 1633 | ACATCAAGAACGCCCTGGAGACTCTCCATGAGAGGTTCTGGGAGATTGTGAGCTG | 1692 | Db | 2713 | CCGAGACCCCTGGTCAACTCTGGAGCTTCTGGAGCTTAACTCCAGTATCATTCGGGAGGTTTC | 2772 |
| Qy | 1657 | GTGACGGCTGCTGAGCTGGAGACTGGAGCTGGAGGAGTACTCTGGGAGATTGTGAGCTG | 1716 | Oy | 2737 | CTGGGAGACCCAGACAGTGTGCCAGGAGGAGCAG 2774 | |
| Db | 1693 | GAATGAGTGTGAGAGTACCCCTGGAGACTTCACTGAACTCATGTCATG | 1752 | Db | 2773 | CTGAGAACCCAGACAGTGTGCCAGGAGT | 2810 |
| Qy | 1717 | CCAGGAGTCCTCAGGCTCTGGCTCTGGAGTTGTGATGATACTACA | 1776 | RESULT 2 | | | |
| Db | 1753 | CCACGGACACATCTCAGCGAGCTGGCTCTGGGCTGGTGTGATGATGAA | 1812 | AL542856 | AL542856 | 935 bp mRNA sequence. | |
| Qy | 1777 | GACTCTCTACCTCATGAACTCAACAGAGATGTTGGAGAGACCCAACTGATT | 1836 | DEFINITION | AL542856 | Homo sapiens PLACENTA | linear EST 12-MAY-2003 |
| Db | 1813 | GACTTCTCAGCTGTCAGGAACTGAGACAGGGATATCGAAACTGAGCTATT | 1872 | ACCESSION | AL542856 | Homo sapiens | CDNA clone CSODE013YE22 |
| Qy | 1837 | GGAGTGGGAGTACAGACAGGGAGATTACTGGTTGCGAGCGCTTGTGAA | 1896 | VERSION | AL542856.2 | EST. | |
| Db | 1873 | ATAGCGAGCATGACAGGGATATTCTACTCTGGTTCTGGCTTGTGAA | 1932 | SOURCE | Homo sapiens | (human) | |
| | | | ORGANISM | Homo sapiens | | | |
| | | | | Eukaryota; Metazoa; Eutheria; Primates; Catarrhini; Hominidae; Homo; | | | |
| | | | | Mammalia; Bacteria; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |

| REFERENCE | 1 (bases 1 to 935) |
|--|--|
| AUTHORS | Li, W.B., Gruber,C., Jessee,J. and Polayes,D. |
| TITLE | Full-length cDNA libraries and normalization |
| JOURNAL | Unpublished |
| COMMENT | On Feb 15, 2001 this sequence version replaced gi:12895311. |
| Contract | Genoscope |
| Genoscope - Centre National de Séquençage | |
| BP 191 91006 EVRY cedex - France | |
| Email | segref@genoscope.cnrs.fr |
| Library | was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1474.f |
| Contact | Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSODE013BC11QPI. |
| FEATURES | source |
| BASE COUNT | 1. .935 |
| ORIGIN | /organism="Homo sapiens" 'db_xref'="Taxon:9606" 'clone'="CSODE013VE22" 'tissue_type'="PLACENTA" 'mol_type'="mRNA" /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. the Not I and EcoRV sites of the pCMVSPORT 6 vector. library was not normalized." |
| BASE COUNT | 229 a 272 c 262 g 172 t |
| RESULT | 3 |
| BM923066 | BM923066 |
| LOCUS | 1081 bp mRNA linear EST 12-MAR-002 |
| DEFINITION | AGENCOURT 6632017 NIH_MSC_118 Homo sapiens cDNA clone IMAGE:5757033 |
| ACCESSION | BM923066.1 |
| VERSION | GI:19373445 |
| KEYWORDS | EST. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens |
| Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| REFERENCE | 5' mRNA sequence. |
| Db | 3477 CTGGGAGGGAGCAAGCAGGAGCTAAATGTGGCGAAGGGGCCAGAGCTAGGGACATGCC 3536 |
| Qy | 40 CTGGTACCGTCGGAATTCCGGATGCGAGGAGGGCCAGCTAGGGACATGCC 99 |
| Db | 3537 AGCCAGGSCCTG3GACACTCTGGTAGAGGAAACACCCCGAACATCAACCTGGAGA 3596 |
| Qy | 100 AGCCGAGGCTGGACACTGTAGAGCGAGTAACACAGCCCACATCAACCTGGAGA 159 |
| Db | 3597 AGACTGGAGGCTGTACACTCTCATGGGTGACACTGTGTCTTACATGTGAGA 3656 |
| Qy | 160 AGACTGGAGGCTGTGACACTCTCATGGGTGACACTGTGTCTTACATGTGAGA 219 |
| Qy | 3657 TCCGGAGGCCACTGGCACCGGATAATGTTAGAGCAACATCCACAGGCCCTGACATCT 3716 |
| Db | 220 TCCGGAGGCCACTGGCACGGATAATGTTAGAGCAACATCCACAGGCCCTGACATCT 279 |
| Qy | 3717 CTCTGAGGGCTCCAAGGGCTTCTGTCACTGCTGGAGCTGAGCAACTGCACTTGCT 3776 |
| Db | 280 CTCTGAGGGCTCCAAGGGCTTCTGTCACTGCTGGAGCTGAGCAACTGCACTTGCT 339 |
| Qy | 3777 GTCAGGGCCAGGGGAAATGTCATGTCAGCTGGAGCTGAGCAACTGCACTTGCT 3836 |
| Db | 340 GTCAGGGCCAGGGGAAATGTCATGTCAGCTGGAGCTGAGCAACTGCACTTGCT 339 |
| Qy | 3837 GAGACACTGGAAAGCTCCTGGAGAGAAAGACTGAAGAAGTGAAGCTGAACTCA 3896 |
| Db | 400 CAGACACTGGAAAGCTCCTGGAGAGAAAGACTGAAGAAGTGAAGCTGAACTCA 459 |
| Qy | 3897 GTCAGGGCCAGGGGAAATGTCATGTCAGCTGGAGCTGAGCAACTGCACTTGCT 3956 |
| Db | 460 GTCAGGGCCAGGGGAAATGTCATGTCAGCTGGAGCTGAGCAACTGCACTTGCT 519 |
| Qy | 3957 GTCAGGGCCAGGGGAAATGTCATGTCAGCTGGAGCTGAGCAACTGCACTTGCT 4016 |
| Db | 520 GTCAGGGCCAGGGGAAATGTCATGTCAGCTGGAGCTGAGCAACTGCACTTGCT 579 |
| Qy | 4017 CTCACCTCTCTCGAGGAGCTGTTCAAGCTGAGGAGCTGAGCACTGAGGAGCTG 4076 |
| Db | 580 CTCACCTCTCTCGAGGAGCTGTTCAAGCTGAGGAGCTGAGCACTGAGGAGCTG 639 |
| FEATURES | source |
| BASE COUNT | 255 a 329 c 286 g 208 t |
| ORIGIN | Best Local Similarity 93.1%; Pred. No. 9; 7e-194; Matches 925; Conservative 0; Mismatches 52; Indels 17; Gaps -9; |
| Qy | 4077 CATGCAGCTCTGGACACATGCTGGAGCAGTGGGGCAAGACTACCTCCACAACTT 4136 |
| Db | 640 CATGCAGCTCTGGACACATGCTGGAGCAGTGGGGCAAGACTACCTCCACAACTT 699 |
| Qy | 4137 CACCCACGGCCAGCCAACTCAAGTGGCAACCTCCACCT 4196 |
| Db | 700 CACCCACGGCCAGCCAACTCAAGTGGCAACCTCCACCT 759 |
| Qy | 4197 GCGGACAGCCGATTCCTCCAGACAGCGCTGAGAGCCCCGAGGTGTACTGGC 4256 |
| Db | 760 GCGGACAGCCGATTCCTCCAGACAGCGCTGAGAGCCCCGAGGTGTACTGGC 819 |
| Qy | 4257 TGTCGCACTGGCAAGGGAGTCGGCTTGTGGGGCATCTGGAGAAGTGTCTGGAG 4316 |
| Db | 820 TGTCGCACTGGCAAGGGAGTCGGCTTGTGGGGCATCTGGAGAAGTGTCTGGAG 879 |
| Qy | 4317 GTCCAGGAGGAGGGCCAGGAGAATCTCCAAATGAGCTGCTGAGCTGTGCCCC 4372 |
| Db | 880 GTCCAGGAGGAGGGCCAGGAGATCTCCAAATGAGCTGCTGAGCTGTGCCCC 935 |

| | | DEFINITION | AGENCOURT 8125945 Lupski_dorsal_root_ganglion Homo sapiens cDNA |
|----------|----------|--|---|
| | | ACCESSION | Clone IMAGE:6177622_5', mRNA sequence. |
| | | VERSION | BU150503 |
| | | KEYWORD | EST. |
| QY | 2648 | ATGCCCTGGACGTCCTGATAGAGAGGCCCCAGACTCTGTCACCTCGTGACTCC | 2707 |
| Db | 16 | ATGCCCTGGACGTCCTGATAGAGAGGCCCCAGACTCTGTCACCTCGTGACTCC | 75 |
| QY | 2708 | TGAAACCCACTATCATGGCGCAGCTGTCTGGAAACCCAGAACAGAACTGCCCAGTC | 2767 |
| Db | 76 | TGAAACCCACTATCATGGCGCAGCTGTCTGGAAACCCAGAACAGAACTGCCCAGTC | 135 |
| QY | 2769 | AGSCCAGGTTCGTGACTGCCTCTGACCTGGGGAGAATCCAAAGGCTAGCCA | 2827 |
| Db | 136 | AGSCCAGGTTCGTGACTGCCTCTGACCTGGGGAGAATCCAAAGGCTAGCCA | 195 |
| QY | 2828 | GCTGTGGAGGCTTCAGCCAGCTTACCGGAGCACATGGGGAGCTGGGGCTAGGCC | 2887 |
| Db | 196 | GCTGTGGAGGCTTCAGCCAGCTTACCGGAGCACATGGGGAGCTGGGGCTAGGCC | 255 |
| QY | 2888 | GCTATGACAGCAGGAGGACTCTCTGTTGCTGCTGAGCTGGAGCTGGGGCTAGGCC | 2947 |
| Db | 256 | GCTATGACAGCAGGAGGACTCTCTGTTGCTGCTGAGCTGGAGCTGGGGCTAGGCC | 315 |
| QY | 2948 | TCTCTGCTGGGGATGGCTTACAGATACGCTCTTGGCCAGACTGATCCACC | 3007 |
| Db | 316 | TCTCTGCTGGGGATGGCTTACAGATACGCTCTTGGCCAGACTGATCCACC | 375 |
| QY | 3008 | CAATCAGAAATTCACTCCAGCTGGCAGAGCCCTTGGACCAATATGCTGACCC | 3067 |
| Db | 376 | CAATCAGAAATTCACTCCAGCTGGCAGAGCCCTTGGACCAATATGCTGACCC | 435 |
| QY | 3068 | TGGAAGCAAACAGAACCTGACCTGAGACAGAGATGGCCATCACCTGTCAC | 3127 |
| Db | 436 | TGGAAGCAAACAGAACCTGACCTGAGACAGAGATGGCCATCACCTGTCAC | 495 |
| QY | 3128 | AGATATGACCCCTTCTGAAACCTCTGAAATGTAATACAGTACCTGACCAAGCG | 3187 |
| Db | 496 | AGATATGACCCCTTCTGAAACCTCTGAAATGTAATACAGTACCTGACCAAGCG | 555 |
| QY | 3188 | CTCTGAGACTGGGCACTGACTCTCTGTACAGAGGGCTTCAATAGTGTTC | 3247 |
| Db | 556 | CTCTGAGACTGGGCACTGACTCTCTGTACAGAGGGCTTCAATAGTGTTC | 615 |
| QY | 3248 | CAACCTCTCCACCAAGCTGGACCAAGGAGAACATCAAGTGTGGCC-CCTCTGGTAC | 3306 |
| Db | 616 | CAACCTCTCCACCAAGCTGGACCAAGGAGAACATCAAGTGTGGCC-CCTCTGGTAC | 675 |
| QY | 3307 | TCTCTGACTACAGCTGGGGCTGACCAACACTCCAGTACCCACATCTGG | 3366 |
| Db | 676 | TCTCTGACTACAGCTGGGGCTGACCAACACTCCAGTACCCACATCTGG | 735 |
| QY | 3367 | AGGGACTCTCTGGAGCTGGGGACTTGGAGACTCACCC-ACACTGTC | 3424 |
| Db | 736 | AGGGACTCTCTGGAGCTGGGGACTTGGAGACTCACCCACACTGTC | 795 |
| QY | 3425 | CTACATCTCTGGAGAGTCACCCCT--ACCTCTCTCTCTACCAAGCACTGG | 3481 |
| Db | 796 | CTACATCTCTGGAGAGTCACCCCTTACCTCTACCAAGCACTGG | 855 |
| QY | 3482 | AGGGACAGCAGGAG-ACTAATGTTG--CAGGGAGGGCAGAGCTGG-ACATCC | 3536 |
| Db | 856 | AGGGACAGCAGGAG-ACTAATGTTG--CAGGGAGGGCAGAGCTGG-ACATCC | 915 |
| QY | 3537 | AGCCAGGCC--TGGACCTGGTAGAGGATGAAA--ACAGCCCCGACATCAACT | 3590 |
| Db | 916 | AGCCAGGCC--TGGACCTGGTAGAGGATGAAA--ACAGCCCCGACATCAACT | 975 |
| QY | 3591 | GAGAAAGACTGGAGCTGTACACTCTT 3624 | 1009 |
| Db | 976 | GAGAAAGACTGGAGCTGTACACTCTT 1009 | 1009 |
| RESULT 4 | BU150503 | 935 bp mRNA linear EST 03-SEP-2002 | |
| LOCUS | | | |

| | | | |
|-----------------------|---|---|------|
| QY | 3784 | GGCCAAAGCGGGAAATGTCATGTCGAGCTCAGAACACTTGCCCTCAGAC | 3843 |
| Db | 301 | GGCCAAAGCGGGAAATGTCATGTCGAGCTCAGAACACTTGCCCTCAGAC | 360 |
| QY | 3844 | TGCPAAGCTCTGTGAGAACGAGAACTGAGAAAGTGAACGACTGGAAC | 3903 |
| Db | 361 | TGCCAAGCTCCCTGAGAGCAGAACGAGAACTGGAAC | 420 |
| QY | 3904 | ATCTCAAGTTCTCTACTGGCACCATACACAGCGACTTGGGTGTCG | 3963 |
| Db | 421 | ATCTCAAGTTCTCTACTGGCACCATACACAGCGACTTGGGTGTCG | 480 |
| QY | 3964 | CAGCCCTTCTCCAAAAGACACTCACCCACTGACAGAGAGGGAACTGAC | 4022 |
| Db | 481 | CAGCCCTTCTCCAAAAGACACTCACCCACTGACAGAGAGGGAACTGAC | 540 |
| QY | 4023 | CTTCTCTCCGAGGACGTTTCACTTCAGACGGGGCATGCCAGATGCC | 4082 |
| Db | 541 | CTTCTCTCCGAGGACGTTTCACTTCAGACGGGGCATGCCAGATGCC | 420 |
| QY | 4083 | ACTCTGGACACATGCTGGACCGTGGCCGAGACTTCACACTCACCA | 4142 |
| Db | 601 | ACTCTGGACACATGCTGGACCGTGGCCGAGACTTCACACTCACCA | 560 |
| QY | 4143 | CGGCCGAGCCAACTCAAGTGCCCTCTCTGAGAGGCCCGAGGCTTAC | 4201 |
| Db | 661 | CGGCCGAGCCAACTCAAGTGCCCTCTCTGAGAGGCCCGAGGCTTAC | 720 |
| QY | 4202 | ACAGCCTATGCCCAAGCCAGCGCTGAGAGGCCCGAGGCTTAC | 4261 |
| Db | 721 | ACAGCCTATGCCCAAGCCAGCGCTGAGAGGCCCGAGGCTTAC | 780 |
| QY | 4262 | CACTGGCA - GCGGAGTGGCTTGTGGGGCATCTGGGACAGT-GGTCTGGAGT | 4318 |
| Db | 781 | CAAGGGCAGGGGAACTCGCCCTGTGTGGGCATCTTCGAAAGGGCTGAGGT | 840 |
| QY | 4319 | GCAGGAGGGTGG -- CGGGAGGGAGATCTCTCAA-TGAGCTGGCACTGTGGCTC | 4374 |
| Db | 841 | GCAGGAGGGTGG -- CGGGAGGGAGATCTCTCAA-TGAGCTGGCACTGTGGCTC | 900 |
| QY | 4375 | TA 4376 | |
| Db | 901 | CA 902 | |
| RESULT 5 | | | |
| BX349938/C | | | |
| DEFINITION | BX349938 Homo sapiens 903 bp mRNA linear EST 05-MAY-2003 | | |
| LOCUS | BX349938 | | |
| VERSION | 1 | | |
| KEYWORDS | EST. | | |
| SOURCE | Homo sapiens (human) | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. | | |
| COMMENT | Contact : Genoscope Genoscope - Centre National de Séquençage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 14744. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0BA10052A12_CS00473_2. | | |
| FEATURES | source /organism="Homo sapiens" | | |
| BASE COUNT | 162 a 242 c 269 g 223 t | | |
| ORIGIN | 7 others | | |
| Query Match | | | |
| Best Local Similarity | 17.4% | | |
| Matches | 833; | | |
| Pred. No. | 1 2e-186; | | |
| Mismatches | 0; | | |
| Indels | 26; | | |
| Gaps | 41; | | |
| QY | 3491 | CTCTGAGAACTTCACCCCTTCTCTGGCTTCTTACAGCACCTGGGGAGC | 3490 |
| Db | 843 | CTAGGACTTAATGGCAGGGCAGANGCTAGNACTTGCCGCCGGCTGGA | 784 |
| QY | 3551 | ACCTGGTAGAGGAATGAAACAGCCCGAGCATCAACCTGGAGAAAGACTGAGCTGG | 3610 |
| Db | 783 | CTCTGGTAGAGGAATGAAACAGCCCGAGCATCAACCTGGAGAAAGACTGAGCTGG | 724 |
| QY | 3611 | TACACTCTTCACTGGGTCAGACGACTGTGTCATTACTGGAGAAATCC----- | 3659 |
| Db | 723 | TACACTCTTCACTGGGTCAGACGACTGTGTCATTACTGGAGAAATCC----- | 664 |
| QY | 3660 | -----GGAGGCCACTTGCCACGGAAATATGTCAG | 3690 |
| Db | 663 | GGAGGCCACTTGCCACGGAAATATGTCAG | 604 |
| QY | 3691 | CACTCCACAGCCCTGAGCTCTCTGAGGGAGCTCCAGGGCTTCTGCAACGTG | 3750 |
| Db | 603 | CACTCCACAGCCCTGAGCTCTCTGAGGGAGCTCCAGGGCTTCTGCAACGTG | 544 |
| QY | 3751 | GTGAGGGTCACTGAGCTGTGCTGAGCTGCTGAGGGCCAGGGAAATGTCATGCTG | 3810 |
| Db | 543 | GTGAGGGTCACTGAGCTGTGCTGAGGGCCAGGGAAATGTCATGCTG | 484 |
| QY | 3811 | GCAGCTCGAGAAACTGACTCTGAGACTCTGGAGGAGCTGGAGGAGAAG | 3870 |
| Db | 483 | GCAGCTCGAGAAACTGACTCTGGAGGAGCTGGAGGAGAAG | 424 |
| QY | 3871 | CTCTGAGAAAGTGTGACTGGAACTCCAGATCTGAGCTTCAGTTCTCTACTGGACCAA | 3930 |
| Db | 423 | CTCTGAGAAAGTGTGACTGGAACTCCAGATCTGAGCTTCAGTTCTCTACTGGACCAA | 364 |
| QY | 3931 | TGACACAGGGAGGACTTGGGTTGGGAGCTTCCAAACCACTCACC | 3990 |
| Db | 363 | TGACACAGGGAGGACTTGGGTTGGGAGCTTCCAAACCACTCACC | 304 |
| QY | 3991 | CACTGAGAGGAGGGACCTGAGCTCACCTCTCTGGAGGAGCTGTTCACTC | 4050 |
| Db | 303 | CACTGAGAGGAGGGACCTGAGCTCACCTCTCTGGAGGAGCTGTTCACTC | 244 |
| QY | 4051 | TGAGACGGGGCATGGGGAGTGGGCACTGGGACTCTGGAGAACACAGCTGGACACAGTGG | 4110 |
| Db | 243 | TGAGACGGGGCATGGGGAGTGGGCACTGGGACTCTGGAGAACACAGCTGG | 184 |
| QY | 4111 | GGCGGAGAGCTCAACACTCACCCAGGGAGCAACTCAAGTGCCCTCT | 4170 |
| Db | 183 | GGCGGAGAGCTCAACACTCACCCAGGGAGCAACTCAAGTGCCCTCT | 124 |
| QY | 4171 | CTGAGAGGCTTACCTCCAGGGAGAGCTGGGAGCTGGGCTGTTGCTG | 4230 |
| Db | 123 | CTGAGAGGCTTACCTCCAGGGAGAGCTGGGAGCTGGGCTGTTGCTG | 64 |

| | | | | | |
|--|----|--|---|--------------------------|---|
| Db | 63 | GAGCCCCGAGGTGCTACTGGCTGCCAG-GSCAGGAGTCGGCCCTGGGG 5 | Db | 574 | CCAAAGGCTTGTGCAAGTGTTGGAGTCATGGAGCTGGCTGTGACCGCCA 515 |
| RESULT | 6 | BX349337/c | DEFINITION | BX349337 | 974 bp mRNA linear EST 05-MAY-2003 |
| LOCUS | | | REFERENCE | BX349337 | Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens |
| DEFINITION | | | AUTHORS | C DNA clone CS0DC023YA08 | 3-PRIME, mRNA sequence. |
| ACCESSION | | | TITLE | BX349337 | |
| VERSION | | | JOURNAL | BX349337.1 | GI:310371419 |
| KEYWORDS | | | COMMENT | | |
| SOURCE | | Homo sapiens (human) | ORGANISM | | |
| ORGANISM | | | | | |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | | |
| REFERENCE | | | 1 (bases 1 to 974) | | |
| AUTHORS | | | Li,W.-B., Gruber, C., Jessee,J. and Polaves,D. | | |
| TITLE | | | Full-length cDNA libraries and normalization | | |
| JOURNAL | | | Unpublished | | |
| COMMENT | | | Contact: Genoscope | | |
| FEATURES | | | Genoscope - Centre National de Sequenage | | |
| source | | | BP_191_91006_EVRY_cedex - France | | |
| Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1474.f, Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0BA10052A12_CS00473_1. | | | | | |
| LOCATION/QUALIFIERS | | | 1. .974 | | |
| /organism="Homo sapiens" | | | /mol_type="mRNA" | | |
| /db_xref="Taxon:9606" | | | /clone="CS0DC023YA08" | | |
| /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED" | | | /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED" | | |
| /notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized." | | | /notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized." | | |
| BASE COUNT | | 177 a 259 c 295 g 241 t 2 others | ORIGIN | | |
| Query Match | | | | | |
| Best Local Similarity | | 16.7% | Score | 732.2; | DB 13; Length 974; |
| Matches | | 91.2% | Pred. No. | 3..2e-179; | |
| | | 849; | Mismatches | 0; | |
| | | | Indels | 35; | |
| | | | Gaps | 47; | |
| | | | | 5; | |
| Oy | | | | | |
| 3409 | | ACTCACACACTGCCAACATTCTGAGAGTCACCCCTACCTCTTGCTCT | LOCUS | BX369782 | 900 bp mRNA linear EST 08-MAY-2003 |
| | | 3468 | DEFINITION | BX369782 | Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens |
| Db | | 928 | ACATCTCCCTCTGCACATT-TGAGAGTTAACCTTCTGCTCT-- | ACCESSION | BX369782 |
| Oy | | 3469 | ACGACACCTGGAGGGACAGCAGGACTAAATGTGGCAGCGAGGAGCTAGG | VERSION | BX369782.1 |
| Db | | 871 | 814 AGCAGCACTGGAGGGACAGCAGGAAATAATGCGC-GAAAGGCCACATCAC | KEYWORDS | GI:30441506 |
| Oy | | 3529 | GACATGCCAGCCAGGCCAGCTGGACTAGGGAAATGAAACAGCCCAACATCAC | SOURCE | |
| Db | | 3589 | 814 GACATGCCAGCCAGGCCAGCTGGACTAGGGAAATGAAACAGCCCAACATCAC | ORGANISM | |
| Oy | | 754 | CTGAGAGAGACTGGAAGCTGTGACACTCTCATGGGTCACTGCTTCTGAGG | REFERENCE | |
| Db | | 755 | 755 CTGAGAGAGACTGGAAGCTGTGACACTCTCATGGGTCACTGCTTCTGAGG | AUTHORS | |
| Oy | | 3649 | 755 TGTGAGAATCC-----GGAGGCCA 3668 | TITLE | |
| Db | | 694 | TGTGAGAATCCGTAGGCCGACCAACCCATGGGAGCTGAGAGGAGGGCCA | JOURNAL | |
| Oy | | 635 | 635 CTGGCCACGGAAATATGTCAGACATCCAAACAGGCCCTGGAGGCT 3648 | COMMENT | |
| Db | | 3669 | 3728 CTGGCCACGGAAATATGTCAGACATCCAAACAGGCCCTGGAGGCT 3728 | FEATURES | |
| Oy | | 634 | 695 CTGGCCACGGAAATATGTCAGACATCCAAACAGGCCCTGGAGGCT 575 | source | |
| Db | | 3729 | 3788 CCCAAGGGCTTCTGTCACCGTGGAGGTCACTGAGCTGGCTAGCTGTACCAAGGGCCA 3788 | | |

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched double-strand cDNA was

| Query | Match | 16.7% | Score | 730.6; | DB | 13; | Length | 900; |
|------------|-----------------------|---|--------------|--------|---|------------|--------|-------------------------------------|
| | Best Local Similarity | 91.1% | Pred. | No. | 7.8e-119; | | | |
| | Matches | 821; | Conservative | 0; | Mismatches | 34; | Indels | 46; |
| QY | 3324 | GGGAGCTGACCAACAACTCCAGTACCTACCCATCTGGAGGGACTCTTGAG | 3383 | | | | | |
| QY | 5 | GGGAGCTGCCCCAACACTNCCTGAGAACCTACCAACTTGNGGGACTCTTGAG | 64 | | | | | |
| Db | 3384 | CATTGGAGGGATGGAACTTGAGCTCACACCAACTGCCAACATCTGAAGAAT | 3443 | | | | | |
| QY | 125 | CAACCCATTACCTCTTGCTCTGCTCTACCAAGACCTTGAGAATT | 184 | | | | | |
| Db | 65 | CATTGGAGGGATGGAACTTGAGCTCACACCAACTGCCAACATCTGAAGAAT | 124 | | | | | |
| QY | 3444 | CAACCCATTACCTCTTGCTCTACCAACTTGAGAACCTTGAGAATT | 3503 | | | | | |
| QY | 3504 | GGCAGCGGGAGGGCCAGAGCTAGGAGATGCCAGGCCAGGGCTGGAGCTGTAG | 3563 | | | | | |
| Db | 185 | GCGAGCGGAAGGGCCAGAGCTAGGAGATGCCAGGCCAGGGCTGGAGCTGTAG | 244 | | | | | |
| QY | 3564 | AATGAAACACGCCGACATCACTTGAGAAGACTGAGCTGGAGAAT | 3623 | | | | | |
| Db | 245 | AATGAAACACGCCGACATCACTTGAGAAGACTGAGCTGGAGAAT | 304 | | | | | |
| QY | 3624 | TGGGTGAAACAGACTTGTCTTACTGTGAGATTC | 3659 | | | | | |
| Db | 305 | TGGGTGAAACAGACTTGTCTTACTGTGAGAATTCGAGGAGCTGAGACTTAATGT | 184 | | | | | |
| QY | 3660 | -----GGAGGCCACTTGCCACGGAAATATGTTCAGCACATCACAGG | 3703 | | | | | |
| Db | 365 | GGGACCTGAGAAGGAGGAGGCCATGGCCACGGAAATATGTTCAGCACATCACAGG | 424 | | | | | |
| QY | 3704 | CCCTGACATCCCTCTGAGGACTCCAAAGSCTTCTGTCACTGCTGGAGTCATGG | 3763 | | | | | |
| Db | 425 | CCCTGACATCCCTCTGAGGACTCCAAAGGCTTCTGTCACTGCTGGAGTCATGG | 484 | | | | | |
| QY | 3764 | ACCTGCTTAGCCGTACCGGGCAAGGGGAATGTCCATCTGGAGCTGAGAACAA | 3823 | | | | | |
| Db | 485 | ACGTGCTTAGCCGTACAGGGCAAGGGGAATGTCCATCTGGAGCTGAGAACAA | 544 | | | | | |
| QY | 3824 | ACTGACTCTGCTCAAGACACTCCAAAGCTCCCTGAGAGAACTGAGAACAGA | 3883 | | | | | |
| Db | 545 | ACTGACTCTGCTCAAGACACTCCAAAGCTCCCTGAGAGAACTGAGAACAGA | 604 | | | | | |
| QY | 3884 | ACTGGAACTCTCAGATGCTCAGTCTCAGTTCTCTTACTTGAGGACAACTACACAGCTG | 3943 | | | | | |
| Db | 605 | ACTGGAACTCTCAGATGCTCAGTTCTCTTACTTGAGGACAACTACACAGCTG | 664 | | | | | |
| QY | 3944 | -AGGAGTTGCGGTGCGTGCAGCTCTTCAAAACACACTCCACCCACTGAGCG | 4002 | | | | | |
| Db | 665 | NAGGAGCTTGCGTNTGCTGCTGAGCTTCTTCAAAACACACTAACCCACTGAGCG | 724 | | | | | |
| QY | 4003 | AGGGGGACACTGACCTCACCTCTCTCCGAGGAGCTGTTGACTTCAGAGCGCG | 4062 | | | | | |
| Db | 725 | AGGGGGACACTGACCTCACCTCTCTCCGAGGAGCTGTTGACTTCAGAGCGCG | 784 | | | | | |
| QY | 4063 | CATGCCAGATGCCATGCCACTCTGGAGAACATGCTGAACTGGCCCAAGACT | 4122 | | | | | |
| Db | 785 | CATGCCAGATGCCATGCCACTCTGGAGAACATGCTGAACTGGCCCAAGACT | 844 | | | | | |
| QY | 4123 | ACCTCCACAACTCCACCAAGGCCAACTCAAGTGCCTCTCTGTAGAGGCCT | 4182 | | | | | |
| Db | 845 | ACTNNCAACACTAC-----CAGAGCCACAACTCACTGCCCCCTCTNCTGAGAGCCT | 899 | | | | | |
| QY | 4183 | T | 4183 | | | | | |
| FEATURES | source | | | | | | | |
| COMMENT | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Cattarrhini; Hominidae; Homo. Email: cgabre-r@mail.nih.gov | | | | | | |
| REFERENCE | | (bases 1 to 861) | | | | | | |
| KEYWORDS | | NID-MGC http://mgc.nci.nih.gov/ | | | | | | |
| TITLE | | National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished | | | | | | |
| ORGANISM | | Homo sapiens | | | | | | |
| ACCESSION | | Build 83160 | | | | | | |
| VERSION | | Build 83160.1 | | | | | | |
| KEYWORD | | EST | | | | | | |
| TITLE | | JOURNAL | | | | | | |
| ORGANISM | | Hom sapiens | | | | | | |
| ACCESSION | | Build 83160 | | | | | | |
| VERSION | | Build 83160.1 | | | | | | |
| KEYWORD | | EST | | | | | | |
| REFERENCE | | Mammalia; Butheria; Primates; Cattarrhini; Hominidae; Homo. Email: cgabre-r@mail.nih.gov | | | | | | |
| AUTHORS | | 1 | | | | | | |
| TITLE | | NID-MGC http://mgc.nci.nih.gov/ | | | | | | |
| ORGANISM | | National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished | | | | | | |
| ACCESSION | | Contact: Robert Strausberg, Ph.D. | | | | | | |
| VERSION | | Email: cgabre-r@mail.nih.gov | | | | | | |
| KEYWORD | | Tissue | | | | | | |
| TITLE | | Preparation: ATCC/PCMV/PRP | | | | | | |
| ORGANISM | | cDNA Library Arrayed: The I.M.A.G.E. Consortium (LLNL) | | | | | | |
| ACCESSION | | DNA Sequencing by: Agencourt Bioscience Corporation | | | | | | |
| VERSION | | Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov | | | | | | |
| KEYWORD | | Plate: LIAM13476 row: 9 column: 07 | | | | | | |
| TITLE | | High quality sequence stop: 641. | | | | | | |
| FEATURES | | Location/Qualifiers | | | | | | |
| BASE COUNT | | 1. . 861 | | | | | | |
| ORIGIN | | /organism="Homo sapiens" | | | | | | |
| FEATURES | | /mol_type="mRNA" | | | | | | |
| BASE COUNT | | /db_xref="taxon:9606" | | | | | | |
| ORIGIN | | /clone IMAGE:6147102" | | | | | | |
| FEATURES | | /tissue="melanotic melanoma" | | | | | | |
| BASE COUNT | | /lab_host="DH10B (phage-resistant)" | | | | | | |
| ORIGIN | | /clone Lib=NID-MGC_72" | | | | | | |
| FEATURES | | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | | | | | | |
| BASE COUNT | | Location/Qualifiers | | | | | | |
| ORIGIN | | 1. . 861 | | | | | | |
| FEATURES | | /organism="Homo sapiens" | | | | | | |
| BASE COUNT | | /mol_type="mRNA" | | | | | | |
| ORIGIN | | /db_xref="taxon:9606" | | | | | | |
| FEATURES | | /clone IMAGE:6147102" | | | | | | |
| BASE COUNT | | /tissue="melanotic melanoma" | | | | | | |
| ORIGIN | | /lab_host="DH10B (phage-resistant)" | | | | | | |
| FEATURES | | /clone Lib=NID-MGC_72" | | | | | | |
| BASE COUNT | | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | | | | | | |
| BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| 215 | a | /organism="Homo sapiens" | 236 | c | /mol_type="mRNA" | 234 | g | /db_xref="taxon:9606" |
| 215 | a | /clone IMAGE:6147102" | 236 | c | /tissue="melanotic melanoma" | 234 | g | /lab_host="DH10B (phage-resistant)" |
| 215 | a | /clone Lib=NID-MGC_72" | 236 | c | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | 234 | g | /clone IMAGE:6147102" |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| 215 | a | /organism="Homo sapiens" | 236 | c | /mol_type="mRNA" | 234 | g | /db_xref="taxon:9606" |
| 215 | a | /clone IMAGE:6147102" | 236 | c | /tissue="melanotic melanoma" | 234 | g | /lab_host="DH10B (phage-resistant)" |
| 215 | a | /clone Lib=NID-MGC_72" | 236 | c | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | 234 | g | /clone IMAGE:6147102" |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| 215 | a | /organism="Homo sapiens" | 236 | c | /mol_type="mRNA" | 234 | g | /db_xref="taxon:9606" |
| 215 | a | /clone IMAGE:6147102" | 236 | c | /tissue="melanotic melanoma" | 234 | g | /lab_host="DH10B (phage-resistant)" |
| 215 | a | /clone Lib=NID-MGC_72" | 236 | c | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | 234 | g | /clone IMAGE:6147102" |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| 215 | a | /organism="Homo sapiens" | 236 | c | /mol_type="mRNA" | 234 | g | /db_xref="taxon:9606" |
| 215 | a | /clone IMAGE:6147102" | 236 | c | /tissue="melanotic melanoma" | 234 | g | /lab_host="DH10B (phage-resistant)" |
| 215 | a | /clone Lib=NID-MGC_72" | 236 | c | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | 234 | g | /clone IMAGE:6147102" |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| 215 | a | /organism="Homo sapiens" | 236 | c | /mol_type="mRNA" | 234 | g | /db_xref="taxon:9606" |
| 215 | a | /clone IMAGE:6147102" | 236 | c | /tissue="melanotic melanoma" | 234 | g | /lab_host="DH10B (phage-resistant)" |
| 215 | a | /clone Lib=NID-MGC_72" | 236 | c | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | 234 | g | /clone IMAGE:6147102" |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| 215 | a | /organism="Homo sapiens" | 236 | c | /mol_type="mRNA" | 234 | g | /db_xref="taxon:9606" |
| 215 | a | /clone IMAGE:6147102" | 236 | c | /tissue="melanotic melanoma" | 234 | g | /lab_host="DH10B (phage-resistant)" |
| 215 | a | /clone Lib=NID-MGC_72" | 236 | c | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | 234 | g | /clone IMAGE:6147102" |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| 215 | a | /organism="Homo sapiens" | 236 | c | /mol_type="mRNA" | 234 | g | /db_xref="taxon:9606" |
| 215 | a | /clone IMAGE:6147102" | 236 | c | /tissue="melanotic melanoma" | 234 | g | /lab_host="DH10B (phage-resistant)" |
| 215 | a | /clone Lib=NID-MGC_72" | 236 | c | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | 234 | g | /clone IMAGE:6147102" |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| 215 | a | /organism="Homo sapiens" | 236 | c | /mol_type="mRNA" | 234 | g | /db_xref="taxon:9606" |
| 215 | a | /clone IMAGE:6147102" | 236 | c | /tissue="melanotic melanoma" | 234 | g | /lab_host="DH10B (phage-resistant)" |
| 215 | a | /clone Lib=NID-MGC_72" | 236 | c | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | 234 | g | /clone IMAGE:6147102" |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| 215 | a | /organism="Homo sapiens" | 236 | c | /mol_type="mRNA" | 234 | g | /db_xref="taxon:9606" |
| 215 | a | /clone IMAGE:6147102" | 236 | c | /tissue="melanotic melanoma" | 234 | g | /lab_host="DH10B (phage-resistant)" |
| 215 | a | /clone Lib=NID-MGC_72" | 236 | c | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | 234 | g | /clone IMAGE:6147102" |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| 215 | a | /organism="Homo sapiens" | 236 | c | /mol_type="mRNA" | 234 | g | /db_xref="taxon:9606" |
| 215 | a | /clone IMAGE:6147102" | 236 | c | /tissue="melanotic melanoma" | 234 | g | /lab_host="DH10B (phage-resistant)" |
| 215 | a | /clone Lib=NID-MGC_72" | 236 | c | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | 234 | g | /clone IMAGE:6147102" |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| 215 | a | /organism="Homo sapiens" | 236 | c | /mol_type="mRNA" | 234 | g | /db_xref="taxon:9606" |
| 215 | a | /clone IMAGE:6147102" | 236 | c | /tissue="melanotic melanoma" | 234 | g | /lab_host="DH10B (phage-resistant)" |
| 215 | a | /clone Lib=NID-MGC_72" | 236 | c | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | 234 | g | /clone IMAGE:6147102" |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| 215 | a | /organism="Homo sapiens" | 236 | c | /mol_type="mRNA" | 234 | g | /db_xref="taxon:9606" |
| 215 | a | /clone IMAGE:6147102" | 236 | c | /tissue="melanotic melanoma" | 234 | g | /lab_host="DH10B (phage-resistant)" |
| 215 | a | /clone Lib=NID-MGC_72" | 236 | c | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies." | 234 | g | /clone IMAGE:6147102" |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES | BASE COUNT | ORIGIN | FEATURES |
| 215 | a | 1. . 861 | 236 | c | 1. . 861 | 234 | g | 1. . 861 |
| 215 | a | /organism="Homo sapiens" | 236 | c | /mol_type="mRNA" | 234 | g | /db_xref="taxon:9606" |
| 215 | a | /clone IMAGE:6147102" | 236 | c | /tissue="melanotic melanoma" | 234 | g | /lab_host="DH10B (phage-resistant)" |
| 215 | a | /clone Lib=NID-MGC_72" | 236 | c | /note=Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT | | | |

QY 1296 GAACTGGCCACCTTACCACTGCGAACATCCTCCGGATTCAACCTTCCTGAA 1355

Db 437 GAACTCGCACCGTACCCACCTGGCGAACATCCTCCGGATTCAACCTTCCTGAA 496

QY 1356 GGCTTCTGTGTTGCATGGAGAAGAACCCAGCTCTTAACAGGCT 1415

Db 497 GGCTTCTGTGTTGCATGGAGAAGAACCCAGCTCTTAACAGGCT 556

QY 1416 GGAGGAGGCCGAAGCTGGAGATCTGGAGATCTACCTTGCTTAACAGGCT 1475

Db 557 GGAGGAGGCCGAAGCTGGAGATCTACCTTGCTTAACAGGCT 616

QY 1476 GAACTGGAGGATACCTTCAGAAGACTGGAGATAAACCGTTAGGGG 1535

Db 617 GAACTGGAGGATACCTTCAGAAGACTGGAGATAAACCGTTAGGGG 676

QY 1536 CAATGACTCTGTGTTGCATGGAGAAGAACCCAGCTCTTAACAGGCT 1595

Db 677 CAATGACTCTGTGTTGCATGGAGAAGAACCCAGCTCTTAACAGGCT 736

QY 1596 CAACATGGAAAGGCCCTGAGCATCTCATGT-GAGTTCTGGCATTTG-TGAC 1653

Db 737 CAACATGGAAAGGCCCTGAGCATCTCATGT-GAGTTCTGGCATTTG-TGAC 796

QY 1654 CTGGTGAACGGT-GTGGAGATCTAACCTGGAGCTTACCAAGGAGAAA 1707

Db 797 CTGGTGAACGGT-GTGGAGCTTACCAACTGGAGCTTACCCGGAGAAA 853

RESULT 9
BI912651

LOCUS BI912651 880 bp mRNA linear EST 16-OCT-2001

DEFINITION 603290988E1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5220333 5,

mRNA sequence.

ACCESSION BI912651

VERSION BI912651.1 GI:16176846

KEYWORDS EST

SOURCE ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 880)

AUTHORS NIH_MGC http://mgc.ncbi.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Straubberg, Ph.D.

Email: rcsrbs@email.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed By: The I.M.A.G.E. Consortium (IIML)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium(IIML) at:

<http://image.IIML.gov>

Place: LUMI553 row: 0 column: 22

High quality sequence start: 27
High quality sequence stop: 726.

FEATURES source

1. .880 /organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="Taxon:9606"

/clone="IMAGE:5220333"

/tissue_type="leukocyte"

/lab_host="DH10B"

/clone_lib="NIH_MGC_118"

/note="Vector: PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source: leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note:

(Invitrogen). Research Genetics tracking code 027. Note: This is a NIH MGC Library."

BASE COUNT 201 a 271 c 228 g 179 t 1 others

Query Match 16.2%; Score 708.2; DB 12; Length 880;

Best Local Similarity 98.3%; Pred. No. 5; 3e-13; Mismatches 8; Indels 5; Gaps 5;

Matches 768; Conservative 0; Mismatches 8; Indels 5; Gaps 5;

DB 2619 TCGAGCCAACTTGTGAGACATCTCCGAATGCGCTTGCTGATAGAGGTC 2678

Db 27 TSGAGCCAACTTGTGAGACATCTCCGAATGCGCTTGAGGTC 86

QY 2679 CAGATCTGGTCAACCTCTGAGACTCTCTGAGCCACTATCATSGGGCAGTGTCT 2738

Db 87 CAGATCTGGTCAACCTCTGAGACTCTGTGAGCTCTGAGCCACTATCATSGGGCAGTGTCT 146

QY 2739 GCGAACCCAGACAAGTCCAGCTGGAGCCAGGCTTGTGAGCTGTCTGAGGTC 2798

Db 147 GCGAACCCAGACAAGTCCAGCTGGAGCCAGGCTTGTGAGCTGTCTGAG 206

QY 2799 CCTGCG-GGAGACTCCCAGAGCTAGCCAGGGTGGAGCCCTGAGCGAGCTACCG 2857

Db 207 CCTGGCGGAGAGACTCCCAGAGCTAGCCAGGGCTTGTGAGGCGCTTACCG 266

QY 2898 GCAGCATGGCGAGCTGGAGCCAGGAGCTTGTGAGGAGACTCTCTGAG 2917

Db 267 GCAGCATGGCGAGCTGGAGGCTAGCCGCTATGAGACGGAGGAGACTCTGTG 326

QY 2918 TCTGCAACCCCTCTCCAGAACATCGCGCTCTGCTGGGGATGGGCTCCAGTA 2977

Db 327 TCTGCAACCCCTCTCCAGAACATCGCGCTCTGCTGGGGATGGGCTCCAGTA 386

QY 2978 CGTCCTCTTGGCCAGACTCATCCACCCAAATCAGAAATTCACCTCCAGCTGCCA 3037

Db 387 CGTCCTCTTGGCCAGACTCATCCACCCAAATCAGAAATTCACCTCCAGCTGCCA 446

QY 3038 GACCCCTTGGACCAATATGCTGAACCAACTTGAAGACAAACAGAACCTGGACCTGA 3097

Db 447 GACCCCTTGGACCAATATGCTGAACCAACTTGAAGACAAACAGAACCTGGACCTGA 506

QY 3098 GACCAACTGGAGATCCAGCTGGCCACTCTGAGATGGCCCTTCTGAGAACCCCTGGGA 3157

Db 507 GACCAACTGGAGATCCAGCTGGCCACTCTGAGATGGCCCTTCTGAGAACCCCTGGGA 566

QY 3158 ATAGTAACATACGGTACCCCATGAAGCCAGATTGAGAACCTGGGAGTGTACTCTGT 3217

Db 567 ATAGTAACATACGGTACCCCATGAAGCCAGATTGAGAACCTGGGAGTGTACTCTGT 626

QY 3218 GTACAGACTGGAGGCTTCAATAGTCTGTCRACC-TCTGTCACCATGGCTGGACCA 3276

Db 687 GTACAGACTGGAGGCTTCAATAGTCTGTCACCATGGCTGGACCA 686

QY 3277 GACATCAAGTGGCCGCCCTGGTCACTCTGACATCA-GCAAGGGAGCTCGAC 3335

Db 687 GACATCAAGTGGCCGCCCTGGTCACTCTGACATCA-GCAAGGGAGCTCGAC 746

QY 3336 AAACAACCTCCAGTGGACCATACCATCTGG-AGGGGACTCTTGTGAGCATGGGG 3393

Db 747 CAGCAACTCCAGTGGACCATCTTGGAGGGAGCTCTGGAGCATGGGG 806

QY 3394 G 3394

Db 807 G 807

RESULT 10
BX354650

LOCUS BX354650 1185 bp mRNA - linear EST 05-MAY-2003

DEFINITION CDNA clone CSUDC033YA08 5'-PRIME, mRNA sequence.

ACCESSION BX354650

QY 3057 GCTTGACCACTGGAGGAAACAGAGACCTGGACAGAGATGCCATCAC 3116
 Db 153 GCTTGAACACTGTGAAACAAACAGAGACCCCTGGACCTGGAGAGATGCCATCAC 212
 QY 3117 CTCCTCCACTCTGAATGAGCCCTCTGAGAACCCCTGGAGATAGTAACAGTACCC 3176
 Db 213 CTGTCGCACTCTGAGAACCCCTGGAGATAGTAACAGTACCC 272
 QY 3177 CTCATCAASCCAGGATTGAGAACCTGGGAGTACTTCTGTGTTACAGGTGAGGTT 3236
 Db 273 CTCATCAASCCAGGATTGAGAACCCCTGGGAGTACTTCTGTGTTACAGGTGAGGTT 332
 QY 3237 CAATAGTGTCCACCTCTGTCACCCAGCTCCACCAACAGAGATGCCATCAC 3296
 Db 333 CAATAGTGTCCACCTCTGTCACCCAGCTCCACCAACAGAGATGCCATCAC 392
 QY 3297 CTCTGGGACTCTCTGACTACAGCAGTGGAGCTCGACCAAAACTCCAGTGACCTACC 3356
 Db 393 CTCTGGGACTCTCTGACTACAGCAGTGGAGCTCGACCAAAACTCCAGTGACCTACC 452
 QY 3357 CTCATCTTGAGGGACTCTCTGAGCATTCGGGAGTGGGACTTGAGBACTCAC 3416
 Db 453 CTCATCTTGAGGGACTCTCTGAGCATTCGGGAGTGGGACTTGAGBACTCAC 512
 QY 3417 CTCACCTGCCAACATTCTGAGAAGTTGACCCATTACTCTCTGCTCTACACAGC 3476
 Db 513 CTCACCTGCCAACATTCTGAGAAGTTGACCCATTACTCTCTGCTCTACACAGC 572
 QY 3477 CTGGGAGGGAGCAGCAGCACTAAATGTTGCACTGGAGGGCAAGCTTAGGACATGCC 3536
 Db 573 CTGGGAGGGAGCAGCAGCACTAAATGTTGCACTGGAGGGCAAGCTTAGGACATGCC 632
 QY 3537 AGCCCAAGGCCTTGAGGACTCTGAGGATG-AAACAGCCGACTTCACCTGG-AG 3594
 Db 633 AGCCCAAGGCCTTGAGGACTCTGAGGATG-AAACAGCCGACTTCACCTGG-AG 692
 QY 3595 AAAGACTGGAGGTGGTACAACCTCTCATGGGGTCAAGGACTGTGTCATTACTGTGAG 3654
 Db 693 AAAGACTGGAGGTGGTACAACCTCTCATGGGGTCAAGGACTGTGTCATTACTGTGAG 752
 QY 3655 ATTCGGGAGGGCCACTTGCCACGGAAATGTTGCACTACAGGCCCCCTGGACATC 3714
 Db 753 ATTCGGGAGGGCCACTTGCCACGGATATGTTGCACTACAGGCCCCCTGGACATC 812
 QY 3715 CTCCTCTGAGGAACTCCCAA-GGCTTGTCAAGT-GGTGAGGTCTGGCT-GGCT 3771
 Db 813 CTCCTCTGAGGAACTCCCAAAGGAGCTTGTCAAGTCTGGGTGGGGATGGCTGGCT 872
 QY 3772 ACCTCTGAGGAACTCCCAAAGGAGGAA 3798
 Db 873 ACCTCTGAGGAACTCCCAAAGGAGGAA 899
 RESULT 12

B1912073 LOCUS BI912073 DEFINITION 60306863|F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5217806 5', mRNA Sequence. ACCESSION B1912073 VERSION B1912073.1 EST: GI:16176045 SOURCE Homo sapiens (human) ORGANISM Homo sapiens (human) Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Primates; Hominoidea; Hominidae; Homo. REFERENCES 1 (bases 1 to 763) NIH_MGC http://mgc.ncbi.nih.gov/ AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC) TITLE Unpublished COMMENT Contact: Robert Strausberg, Ph.D. Email: craigs@email.nih.gov Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL) DNA Sequencing by: Invitae Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILNL at: http://image.llnl.gov Plate: LILNL1546 Row: n column: 07 High quality sequence stop: 758. Location/Qualifiers 1. .763 FEATURES source /organism="Homo sapiens" /mol_type="mRNA" /db_xref="ttaxon:9606" /clone="IMAGE:5217606" /tissue_type="leukocyte" /lab_host="DH10B" /clone_lbs="NIH MGC 118" /note="vector: PCMV-SPOR6; site 1: NotI; site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dt primed and directionally cloned (EcoRV site is destroyed upon Cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH MGC library." BASE COUNT 177 a 238 c 193 g 155 t ORIGIN Query Match 15.3% Score 671.4 DB 12 Length 763: Best Local Similarity 95.9%; Pred. No. 1. 9e-163; Matches 732; Conservative 0; Mismatches 26; Indels 5; Gap 4; DB 2 TGTATTCGCACTTGTCACCTTGTCATCTGCTGAGCTCCATAGAG 60
 QY 2672 AGGTGCCAGAGTCCTGTCACACTCTGGACTCTCTGGACCTCTGGCATAGAG 2671
 Db 61 AGGTGCCAGAGTCCTGTCACACTTGTCATCTGCTGAGCTCCATAGAG 60
 QY 2732 TGTTCCCTGGAAACCCGGAGCAAGTGGCCAGGGCCAGGGCTGGAGGCTTGCTAATCTGG 2791
 Db 121 TGTTCCTGGAAACCCGGAGCAAGTGGCCAGGGCTGGAGGCTTGCTAATCTGG 179
 QY 2792 TTCTGACCTGGGGAGACTCCAAAGGCTGGCCAGGTGGAGGCTTGCTAATCTGGACCT 2851
 Db 180 TTCTGACCTGGGGAGACTCCAAAGGCTGGAGGCTTGCTAATCTGGACCT 239
 QY 2852 ACCGGAGGAGCATGGGGAGCTGGGGTCAAGGCCCATACTGACACGGAGGACTCT 2911
 Db 240 ACCGGAGGAGCATGGGGAGCTGGGGTCAAGGCCCATACTGACACGGAGGACTCT 299
 QY 2912 CTCTGGCTGAGCCTTCTTCAGAACATCCAGCTCTCTCTGGGGATGGCTCC 2971
 Db 300 CTCTGGCTGAGCCTTCTTCAGAACATCCAGCTCTCTGGGGATGGCTCC 359
 QY 2972 CAGATAGCTCCCTCTGGCCCGAGCTGCACTCCAAATGAGAAATTCACCTCCAGC 3031
 Db 360 CAGATAGCTCCCTCTGGCCCGAGCTGCACTCCAGC 419
 QY 3032 TGGCCAGGCCCTTGGACCAA-TATGCTGACCAAGAACAGAGACCT 3089
 Db 420 TGGCCAGGCCCTTGGACCAAATATGCTGACCAAGAACAGAGACCT 479
 QY 3090 GCACTCTGGAGAGATGTCCTCCACTCCAGCTGAGATGAGCCCTCTGTGAGAC 3149
 Db 480 GCACTCTGGAGAGATGTCCTCCACTCCAGCTGAGAC 539
 QY 3150 CCTCTGGAGATGACTACAGTACCTGAGGAGGAGCTGGGAGTGGAGCTGGAGTGGAG 3209
 Db 540 CCCTCGAGATGACTACAGTACCTGAGGAGCTGGGAGTGGAGCTGGAGTGGAG 599

QY 3210 CTCCTGTGAGACTGGAGGCTCCATAGTGTCCAC-CTCTGTCCACCAAGCTCC 3268
Db 600 CTCTGTGAGACTGGAGGCTCCATAGTGTCCACCAAGCTCC 659

QY 3269 GACAGCAGACCATAAAGTGAGCTGCCCTGGTGACTCTCTGACTACAGCTGGGG 3328
Db 660 GACAGCAGACCATAAAGTGAGCTGCCCTGGTGACTCTCTGACTACAGCTGGGG 719

QY 3329 CTGACCAACAACTCCAGTGCCTACCCACATCTTGAGGGG 3371
Db 720 CTGACCAACAACTCCAGTGCCTACCCACATCTTGAGGGG 762

RESULT 13

LOCUS BI488437 823 bp mRNA linear EST 28-AUG-2001
DEFINITION 603020936F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5191712 5',
mRNA sequence.

ACCESSION BI488437
VERSION BI488437.1 GI:15327665
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 823)
AUTHORS NIH-MGC http://mgc.ncbi.nlm.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIHL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIHL at:
<http://image.lnl.gov>
Plate: LLM1479 row: 9 column: 09
High quality sequence stop: 808.

FEATURES source
Location/Qualifiers

1. .823 /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5191712"
/lab_host="DH10B"
/clone_lib="NIH MGC_114"
/note="Organ: bfin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (Destroyed); RNA source anonymous pool of 6
male brains, age range 23-79 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH MGC Library."
BASE COUNT 213 a 239 c 214 g 157 t
ORIGIN

Query Match 14.3%; Score 627.9; DB 12; length 823;
Best Local Similarity 97.2%; Pred. No. 4.7e-152;
Matches 660; Conservative 0; Mismatches 17; Indels 2; Gaps 2;

QY 3057 GTCTGAACCACTTGGAGAACAGAGACCTGGACCTGGAGATGCCATAC 3116
Db 146 GTCTGAACCACTTGGAGAACAGAGACCTGGACCTGGAGATGCCATAC 205

QY 3117 CTGRCGCACTCGATGAGCCCTCTGAGAACCCCTGGAGATGAACTACAGCTGACCC 3176
Db 206 CTGTCGCACTCGATGAGCCCTCTGAGAACCCCTGGAGATGAACTACAGCTGACCC 265

3177 CATGAAAGCCAGCCATTGAGAACCTGGGCACTTCTGTTACAGCTGAGAGGCTC 3236

QY 265 CATCAAGCAGCCATTGAGAACCTGGGCACTGAGCTGGAGGCTC 325
Db 3237 CAATAGTGTCCACCTCTGTCACCACTCCGACCCAGACATCACAGTGTTGGCC 3296
QY 326 CAATAGTGTCCACCTCTGTCACCACTCCGACCCAGACATCACAGTGTTGGCC 385
Db 3297 CCTGGGTACTCTGACTACAGCAGTGACCTCTGTCACCACTCCGACCCAGACATCACAGTGTTGGCC 3356
QY 336 CCTGGGTACTCTGACTACAGCAGTGACCTCTGTCACCACTCCGACCCAGACATCACAGTGTTGGCC 445
Db 3357 CACATCTGGAGGGACTCTCTGGACATGGAGSGATSGAAGCTTGAGACTCAC 3416
QY 446 CACATCTGGAGGGACTCTCTGGACATGGAGSGATSGAAGCTTGAGACTCAC 505
Db 3417 CACACTGCCAACATCTGAGAGTTCACCTTACCTCCCTGGCTCTCACGAC 3476
QY 506 CACACTGCCAACATCTGAGAGTTCACCTTACCTCCCTGGCTCTCACGAC 565
Db 3477 CTGGGAGGGACAGCAGGACTAAATGTCGACGGAGCTCAACCTTACCTCC 3536
QY 556 CTGGGAGGGACAGCAGGACTAAATGTCGACGGAGCTCAACCTTACCTCC 625

Db 3537 AGGCCAGCCTGGAGCTCTGGAGCTGGAGCAATGAAAAGAGCCGACATCACTGGAGA 3596
QY 626 AGGCCAGCCTGGAGCTGGAGCAATGAAAAGAGCCGACATCACTGGAGA 685
Db 3597 AGACTGGAGCTGTCAACTCTTCATGGGTCAACGACTGTGTCATRACTGAGA 3655
QY 636 AGACTGGAGCTGTCAACTCTTCATGGGTCAACGACTGTGTCATRACTGAGA 745
Db 3656 ATCCGAGGCCACTGGCACCGAATATGTCAGCAATCCACAGCCCCTGGACATCC 3715
QY 716 ATCCGAGGCCACTGGCACCGAATATGTCAGCAATCCACAGCCCCTGGACATCC 804
Db 3716 TCTCTGAGGAGCTCCAAAG 3734
QY 805 TCTCTGAGGAGCTCCAAAG 823

RESULT 14

LOCUS BM714799 607 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-EJO-abs-f-21-0-UI.r1 UI-E-EJO Homo sapiens cDNA clone
ACCESSION BM714799
VERSION BM714799.1 GI:19028057
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 607)
AUTHORS Ronald,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
PUBLISHED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bentoo-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq Primer: M13 Reverse.

| FEATURES Source | |
|---|--|
| location/Qualifiers | |
| 1. .607 /organism="Homo sapiens" | |
| /mol_type="mRNA" | |
| /db_xref="axon:9606" | |
| /clone="UI-E-EJ0-ahs-f-21-0-UI" | |
| /tissue_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid" | |
| /dev_stage="fetal and adult" | |
| /lab_host=DH10B (Life Technologies) (T1 phage resistant)" | |
| /clone lib="UI-E-EJ0" | |
| /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-EJ0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. Sequence tags for this library are: fetal eyes, AGATCCAGA ; lens, CGTTAGCA, eye anterior segment, AACCCGAT; optic nerve, CCAATAGG; retina, CCACG; Retina Foveal and Macular, GRCG; and Choroid, ACTCA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)." | |
| BASE COUNT | 156 a 178 c 156 g 117 t |
| ORIGIN | |
| Query Match | 13.9%; Score 607; DB 12; Length 607; |
| Best Local Similarity | 100.0%; Pred. No. 9.9e-147; Mismatches 0; Indels 0; Gaps 0; |
| Matches | 607; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| Query | 3529 GACATGCCAGCCAGGCCTGGACCTGTAGCGGAATGAAACAGCCCCACATAC 3588 |
| Db | 1 GACATGCCAGCCAGGCCTGGACCTGTAGCGGAATGAAACAGCCCCACATAC 60 |
| Query | 3589 CTGGAGAAAGACTGGAGCTGGCACACTCTCATGGGGTCAACGACTTGTCATTC 3648 |
| Db | 61 CTGGAGAAAGACTGGAGCTGGCACACTCTCATGGGTCAACGACTTGTCATTC 120 |
| Query | 3649 TGTGAGAATCCGAGGCCACTTGCCAGGAATATGTCAGCACACATCAAAGGCCCTG 3708 |
| Db | 121 TGTGAGAATCCGAGGCCACTTGCCAGGAATATGTCAGCACACATCAAAGGCCCTG 180 |
| Query | 3709 GACATCCCTCTGGAGGACTCCAGGGCTTGTCAACGCTGGAGCTGGAGCTG 3768 |
| Db | 181 GACATCCCTCTGGAGGACTCCAGGGCTTGTCAACGCTGGAGCTG 240 |
| Query | 3769 GCTGAGCCGTTACCGAGGGCAAGCGGGAAATGTGCCATGCTGCCAGTCAGACAATCG 3828 |
| Db | 241 GCTGAGCCGTTACCGAGGGCAAGCGGGAAATGTGCCATGCTGCCAGTCAGACAATCG 300 |
| Query | 3829 ACTTGCCCTCAGAACACTCCAAACTCTGGAGAACAGAACGAAAGTGAAAGTGACTGG 3888 |
| Db | 301 ACTTGCCCTCAGAACACTCCAAACTCTGGAGAACAGAACGAAAGTGACTGG 360 |
| Query | 3889 AACCTCCAGCATGGCATTCGGCTCTCTGGACCAATACACACAGCTGGAC 3948 |
| Db | 361 AACCTCCAGCATGGCATTCGGCTCTGGACCAATACACACAGCTGGAC 420 |
| Query | 3949 TTGCGGGTGTGGTCAACAAACACACTCACCCACTGACCGAGAGGG 4008 |
| Db | 421 TTGCGGGTGTGGTCAACAAACACACTCACCCACTGACCGAGAGGG 480 |
| Query | 4009 GACACTGACCTCACTTCTCGAGGACTTTCTCTGAGGCCGAGCTGCC 4068 |
| Db | 481 GACACTGACCTCACTTCTCGAGGACTTTCTCTGAGGCCGAGCTGCC 540 |
| Query | 4069 GAGATGCCATCCACTCTGGACAAACATGCTGAAACCAGTGGCCGAAGACTACCTCC 4128 |
| RESULT 15 | |
| AX015993 | |
| LOCUS | AK015993 |
| DEFINITION | Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone: 493539A06 product:weakly similar to PHOSPHOLIPASE B (Rattus norvegicus), full insert sequence. |
| ACCESSION | AK015993 |
| VERSION | AK015993.1 |
| KEYWORDS | HTC; CAP trapper. |
| SOURCE | Mus musculus (house mouse) |
| ORGANISM | Mus musculus |
| REFERENCE | |
| AUTHORS | Carninci, P. and Hayashizaki, Y. |
| TITLE | High-efficiency full-length cDNA cloning |
| JOURNAL | Meth. Enzymol., 303, 19-44 (1999) |
| PUBMED | 9349636 |
| REFERENCE | |
| AUTHORS | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. |
| TITLE | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes |
| JOURNAL | Genome Res., 10 (10), 1617-1630 (2000) |
| PUBMED | 11042159 |
| REFERENCE | |
| AUTHORS | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watsukiki, M., Yoneda, Y., Isikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawaji, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. |
| TITLE | RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer |
| JOURNAL | Genome Res., 10 (11), 1757-1771 (2000) |
| PUBMED | 10520913 |
| REFERENCE | |
| AUTHORS | Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kyosava, H., Kondo, S., Yamana, K., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadori, K., Matsuda, H., Ashburner, M., Batyalov, S., Cesavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaiko, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stabili, P., Supek, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blike, J., Boiffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Garfield, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamira, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, J., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshiida, K., Habegawa, Y., Kawaji, H., Kohetsuki, S., and Hayashizaki, Y. |
| TITLE | Functional annotation of a full-length mouse cDNA collection |
| JOURNAL | Nature, 409 (6821), 685-690 (2001) |
| PUBMED | 11217851 |
| REFERENCE | |

| AUTHORS | The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. |
|--|---|
| TITLE | Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length DNAs |
| JOURNAL | Nature 420, 563-573 (2002) |
| REFERENCE | 6 (bases 1 to 1562) |
| ADACHI, J., AIZAWA, K., AKAHIRA, S., AKIMURA, T., ARAI, A., AONO, H., ARAKAWA, T., BONO, H., CARNINI, P., FUKUDA, S., FUKUNISHI, Y., FURUNO, M., HANAGAKI, T., HARAI, A., HAYATSU, N., HIRAMOTO, K., HIRAKA, T., HORI, F., IMOTANI, K., ISHII, Y., ITOH, M., IWASA, M., KASUBAWA, T., KATO, H., KAWAI, J., KOIMA, Y., KONNO, H., KOURADA, M., KOYA, S., KURIHARA, C., MATSUYAMA, T., MIYAZAKI, A., NISHI, K., NOMURA, K., NUMAZAKI, R., OHNO, M., OKADA, Y., OKIDO, T., OWA, C., SAITO, H., SAITO, R., SAKAI, C., SAKAI, K., SANO, H., SASAKI, D., SUZUKI, H., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TANAKA, T., TEIJIMA, Y., TOYA, T., YAMAMURA, T., YESUNISHI, A., YOSHIDA, K., YOSHIMURA, M., MURAMATSU, M., and HAYASHIZAKI, Y. | |
| TITLE | Direct Submission |
| JOURNAL | Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 220-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) |
| COMMENT | Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. |
| CDNA library | CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGTCAGTTAACTTAACTTACCCCCCCCC 3'], cDNA was prepared by using triethylsioe thermo-activated reverse transcriptase and subsequently enriched for full length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGTCAGTTAACTTAACTTACCCCCCCCC 3']. cDNA was cleaved with BamHI and XbaI. Vector: a modified pBluecript KS(+) after bulk excision from lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B. |
| FEATURES | location/Qualifiers |
| source | 1. .1562 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:4930539A06" /db_xref="NCI:1008469" /db_xref="taxon:10090" /clone="4930539A06" /sex="male" /tissue_type="testis" /clone Lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 192. .1376 /note="unnamed protein product; putative weakly similar to PHOSPHOLIPASE B [Rattus norvegicus] (SPTR 054728, evidence: FASTY, 84%ID, 26.6%length, match=1158)" /codon_start=1 /protein_id="BAB30072.1" /db_xref="GI:12854561" /db_xref="NCI:1022406" /translation="MEIIPGSPVGLLULLIGQGSQIHGGSGENTLAWQSQQFWT LKDPFRCCKPKLLELSLSESHSLRSRDIKLVIAQNPERPLASSTINNEKQSI KNOPQDVGGMGIVMVLISIURHPSVUMPTSPKGPTAVHPTAEDWIQAKELVRLK DRFDLDEBKDWKLITVIFNSNTSQCQSKSHMRMHMLWGDYLHVEVRAF VNLVDLSBVLMANDLDRQBTGSPAPPEVKCTTETTVAWMWSQBAEWDLASSKF NHEHTEAVWFOFPDELEPLERSPQDPTILARLWNMSMEPVQGDGLNTARKT MKCPSEBSERSPYLPFTYKNSYQARLKPITKLQWKGSEBFCDPKNPNSIPTGTMSC SV" SY" |

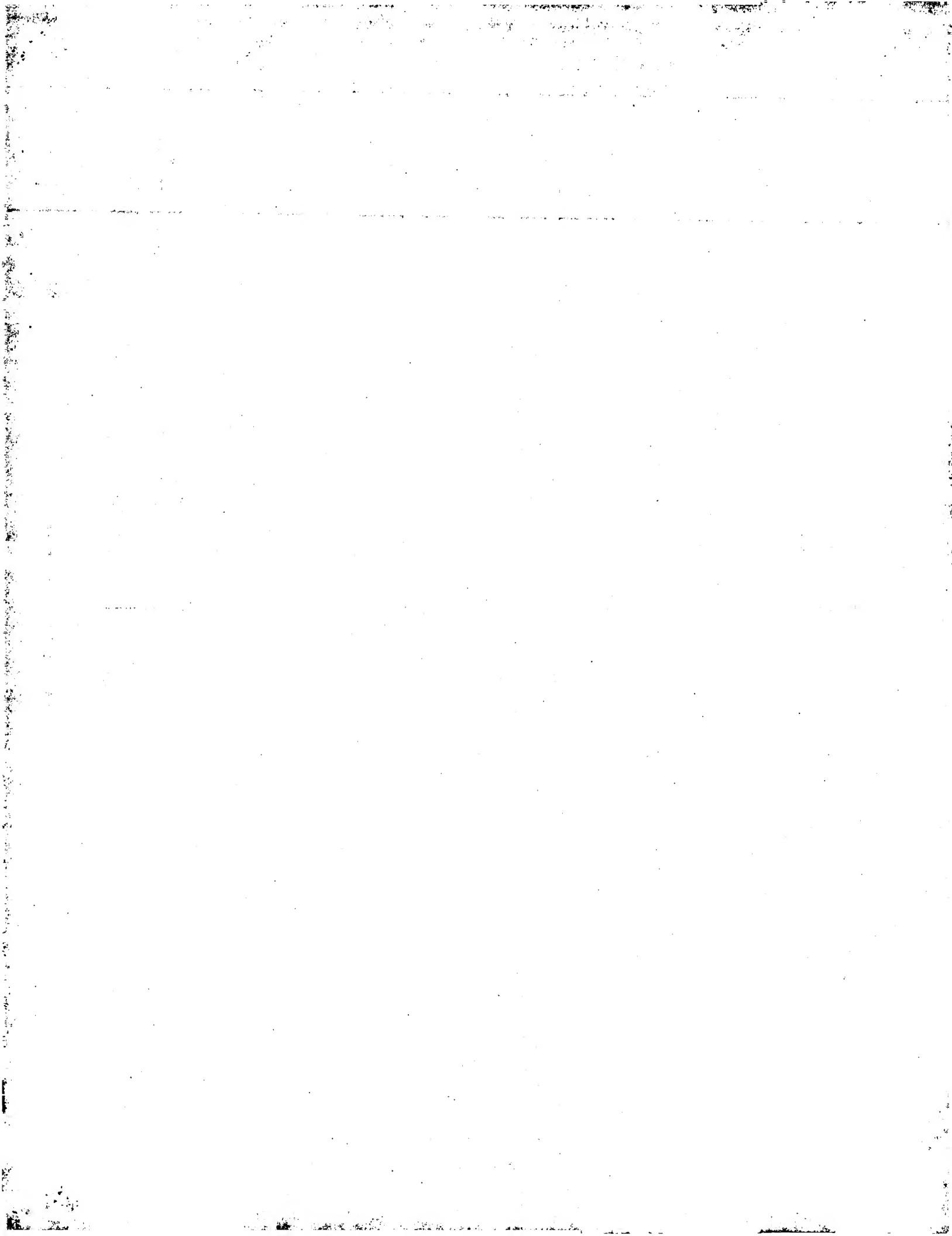
| ORIGIN | Query Match Best Local Similarity 13.2%; Score 578; DB 11; Length 1562; Matches 822; Conservative 0; Mismatches 300; Indels 24; Gaps 4; |
|--------|---|
| QY | 28 CTGGAGCTCTGTCTCTGGCAAGGGACCCCTCGAGTCATACCTCTCTAGAAAG 87 |
| Do | 225 CTGCTGCTGCTACTCTGCTGGACAGGCCCTCCAATCATGCTCTGAGAG 284 |
| QY | 88 AGTACATGGAAAGGCACATAAGGCCAG-----AGACCTGAGAATCTCAGATCCA 141 |
| Do | 285 AACCATGGCATGGCACTGGCAAGCTGGTTGGACCTGAGAATTCTCCAGACCCGAGC 344 |
| QY | 405 TCAGACATTAACCTGTCGCAACCCAGAAATTCTCTGAGCT 464 |
| QY | 262 GGGATCTGGAGAG-----CAAAGCTGACTGAAGAACGGCACAGGGTGATG 312 |
| Do | 465 GGACGATCACTGGAAACTCAAGCTAACAGCAACACAGGATATGCTAG 524 |
| QY | 313 GGAGTGGAGACACTCCTTCAGACATCATCACATGAGCTTTCAGTCCTGTCATCT 372 |
| Do | 585 ACGGTTCTCTGGAGAGGACTGCACTACTGCTGAGATTGTTGAGATTCAG 644 |
| Do | 525 GGATCATGACAGTCTTCAGATATCATCAGACATTCAGCTGCTCTGTCATGCC 584 |
| QY | 373 GTCGCCCCACTCGAAGAGTCATACCCAGATGTGGTGAGAAGACTGGATCAG 432 |
| Do | 615 GCTAAGGCTGGTGGAGACACCCGCAACTTGACTTGAGTTGAGAGACTGG 704 |
| QY | 493 AACCTCATCATGTTCTCAAGTAATGCAACCCAGCTGTTACCTGTCGCCCCTGTC 552 |
| Do | 705 AACTCATCATGTTCTCAACACAAGCCAGTCACCTGTCGTCCTGTCAG 764 |
| QY | 553 CAGATGGCTTCGGGGGGCGCTGGATGGGGGGTGCTGAGACTACCTGGAG 612 |
| Do | 765 CAGAACTCACTGTAGTCAGGCACATGGAGTCGTTGGGGGTGCTGAGATTCAG 824 |
| QY | 613 CAGGAGTCCCAGAGCTTGGATTCAGGTCATGGAGTCGTTGGGGGTGCTGAGATTC 672 |
| Do | 825 CATGGGGCCAGAGCTTGGATTCAGGTCATGGAGTCGTTGGGGGTGCTGAGATTC 884 |
| QY | 673 CGTCAGTACCGCACTGGCTCAGCCCTGTAATGTCAGAGGG 732 |
| Do | 885 CTCAGCATCACAGAGCTGGCTTCAGCCCTGCAACAGAGGGTTGCAATGACA--GAG 941 |
| QY | 733 ACCACCGGCTGCCAAGGGTGTGATGCTGGTCTATCAGGAACCTGGACACCTTC 792 |
| Do | 942 ACACAGGACTATCCAGACTGTCATGAGTCAGTGGCTTCTAGGAACGCTTCTTC 1001 |
| QY | 793 CTGGCTTCAGAGTAGTCAGCATGAGTCAGTGGCTTCAGGTTTCAGCTTCTC 852 |
| Do | 1002 CTGGCTTCAGAGTAGTCAGCATGAGTCAGTGGCTTCAGGTTTCAGCTTCTC 1061 |
| QY | 833 TATGAGACACCCATCTCACCTCGAGGACCCCGACTCCAGGATCTACCTGCTG 912 |
| Do | 1062 GATGAATGAACT----ACCTGAAAGTGCCTGGCCCTCAGGAGGCCCTATGTCTCAGTC 1115 |
| QY | 913 GCTCTGGCTCTCTGGAGTAGGTGAGCTGGAGCAAGGGAGAGATGAGCTG 972 |
| Do | 1116 GCTCTGGCTCTCTGGAGTAGCATGATGGACCTGGTCAGAGGATGGCTACTCAT 1175 |
| QY | 973 GTAAACACGGGCAATGAGTGGCTCTCAGGAGGCCCTATGTCTCAGTC 1032 |
| Do | 1176 ACTCGAGGAGAAGACHATGAATGCTCCCTCTGAGGAGAGTGCCTATCTGTCACCTAC 1235 |

BASE COUNT

441 a 386 c 380 g 355 t

| | | | |
|----|------|---|------|
| QY | 1033 | AGAACAGCAACTACCTGACCGACTGCAGAACCCAAAGACAGCTGAGGTAGAGAA | 1092 |
| Db | 1236 | AATAATAGCAATACCACTTACCGCCAGACAGCTATAACAAAGCTTCAGATGAAGAA | 1295 |
| QY | 1093 | GGAGGGAAATCAGATGTCCTGACAAGAACCCCTCCGATACGGTTCCCACCTCAGTTCAT | 1152 |
| Db | 1296 | GATCAGAATTACCTGTCCTGACAGAACCCCTCGAACTCAATCCCCAACAGGTACA | 1355 |
| QY | 1153 | AGGCTG | 1158 |
| Db | 1356 | TGGCTG | 1361 |

Search completed: January 6, 2004, 14:41:56
Job time : 8450 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2004, 18:56:30 ; Search time 27 Seconds

(without alignments)
2539.442 Million cell updates/sec

Title: US-10-054-691-2

Perfect score: 7766

Sequence: 1 MGURPGIFULELLLILGGQT..... RCRRGGRRADPPMSLRTVAL 1458

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_41;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|--------|--------------------|-------|-------------|
| 1 | 5987.5 | 77.1 | 158 | PHIX_RABBIT |
| 2 | 149 | 1.9 | 355 | YAMB_SCHPO |
| 3 | 138 | 1.8 | 1455 | FACR_HUMAN |
| 4 | 131.5 | 1.7 | 1361 | GIL4_XENIA |
| 5 | 128 | 1.6 | 1155 | C2TA_MOUSE |
| 6 | 126.5 | 1.6 | 1534 | YFBS_ECOLI |
| 7 | 126.5 | 1.6 | 6885 | SNED_HUMAN |
| 8 | 125 | 1.6 | 928 | ATCL_YARII |
| 9 | 125 | 1.6 | 1178 | MNDA YEAST |
| 10 | 123.5 | 1.6 | 1534 | YFAS_ECOLI |
| 11 | 123 | 1.6 | 1 | GCN1_YEAST |
| 12 | 123 | 1.6 | 1039 | PDP2_ARATH |
| 13 | 122.5 | 1.6 | 7113 | GALK_SCHPO |
| 14 | 122 | 1.6 | 1111 | SCT1_SCHPO |
| 15 | 121.5 | 1.6 | 1997 | TOLI_DROME |
| 16 | 121.5 | 1.6 | 8197 | SNEL_HUMAN |
| 17 | 121 | 1.6 | 953 | ATCL_BOVIN |
| 18 | 121 | 1.6 | 4725 | DYIC_DIDI |
| 19 | 120.5 | 1.6 | 2143 | APC_HUMAN |
| 20 | 119.5 | 1.5 | 1192 | RTH4_HUMAN |
| 21 | 119.5 | 1.5 | 1276 | BXO_GLOBO |
| 22 | 117.5 | 1.5 | 1279 | BSC7_PSEFL |
| 23 | 116.5 | 1.5 | 1302 | UBPY_CABEL |
| 24 | 116.5 | 1.5 | 2167 | BEM2_YEAST |
| 25 | 116.5 | 1.5 | 2169 | TEG9_HSVA |
| 26 | 116.5 | 1.5 | 919 | ATCL_HUMAN |
| 27 | 116 | 1.5 | 1420 | SRB9_YEAST |
| 28 | 116 | 1.5 | 1589 | CC5_YEAST |
| 29 | 115.5 | 1.5 | 1 | G6PI_BUCAP |
| 30 | 115.5 | 1.5 | 1449 | BIG3_HUMAN |
| 31 | 115.5 | 1.5 | 5596 | MDNI_HUMAN |
| 32 | 115.5 | 1.5 | 1301 | SAC1_YEAST |
| 33 | 114.5 | 1.5 | 1498 | Y1M9_CLOAB |

ALIGNMENTS

| | | | | |
|---|--|-------------------------------------|------|----------|
| RESULT 1 | PHIX_RABBIT | STANDARD; | PRT; | 1458 AA. |
| ID | PHIX_RABBIT | | | |
| AC | Q05017; | | | |
| DT | 01-FEB-1994 | (Rel. 28, Created) | | |
| DT | 01-FEB-1994 | (Rel. 28, Last sequence update) | | |
| DT | 01-NOV-1997 | (Rel. 35, Last annotation update) | | |
| DB | Phospholipase A2RAB-B precursor (EC 3.1.-.-). | | | |
| OS | Oryctolagus cuniculus (Rabbit). | | | |
| OC | Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus. | | | |
| OX | NCBI_TaxID=9986; | | | |
| RN | SEQUENCE FROM N.A., AND CHARACTERIZATION. | | | |
| RC | TISSUE=Intestine; MEDLINE=9286118; PubMed=8509424; | | | |
| RX | Boll W., Schmid-Chanda T., Semenza G., Mantei N.; "Messenger RNAs expressed in intestine of adult but not baby rabbits: isolation of cognate cDNAs and characterization of a novel border protein with esterase and phospholipase activity." J. Biol. Chem. 268:2901-2911(1993). | | | |
| RU | -I- FUNCTION: HAS ESTERASE AND PHOSPHOLIPASE A2-LYSOPHOSPHOLIPASE ACTIVITY. CAN CONVERT PHOSPHATIDYLCHOLINE TO FATTY ACIDS AND GLYCEROPHOSPHOCHOLINE. COULD BE INVOLVED IN UPTAKE OF DIETARY LIPIDS POSSIBLY INCLUDING LONG CHAIN RETINYL ESTERS. | | | |
| CC | -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. BRUSH BORDER. | | | |
| CC | -I- TISSUE SPECIFICITY: INTESTINE. | | | |
| CC | -I- DEVELOPMENTAL STAGE: EXPRESSED IN THE INTESTINE OF ADULT BUT NOT BABY RABBITS. | | | |
| CC | -I- SIMILARITY: BELONGS TO THE "GDSL" FAMILY OF LIPOLYTIC ENZYMES. | | | |
| This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is not modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_licensee@isb-sib.ch). | | | | |
| EMBL; Z12841; CAJ78103.1; -. | | | | |
| DR | PIR; A45665; A45665. | | | |
| DR | InterPro; IPR001087; Lipase_GDSL. | | | |
| DR | Pfam; PF0057; Lipase_GDSL. | | | |
| DR | PROSITE; PS01098; LIPASE_GDSL_SER; 2. | | | |
| KW | Hydrolease; Repeat; Signal; Transmembrane. | | | |
| FT | 1 | POTENTIAL. | | |
| FT | 1 | PHOSPHOLIPASE A2RAB-B. | | |
| FT | 20 | EXTRACELLULAR (POTENTIAL). | | |
| FT | 20 | TRANSMEM | 1416 | 1419 |
| FT | 1419 | POTENTIAL. | | |
| FT | 1440 | CYTOSPLASMIC (POTENTIAL). | | |
| FT | 1458 | 4 X 308-326 AA APPROXIMATE REPEATS. | | |
| FT | 1403 | O64566 rattus norviger | | |
| FT | 1403 | O9erb4 ratti norvegicus | | |
| FT | 1403 | O12955 homo sapiens | | |
| FT | 1403 | O92973 homo sapiens | | |
| FT | 1403 | P43565 saccharomyces cerevisiae | | |
| FT | 1403 | P75377 mycoplasma pneumoniae | | |
| FT | 1403 | Q13474 homo sapiens | | |
| FT | 1403 | Q14188 schizosaccharomyces pombe | | |
| FT | 1403 | Q9byv6 homo sapiens | | |
| FT | 1403 | Q59722 schizosaccharomyces pombe | | |
| FT | 1403 | Q09933 schizosaccharomyces pombe | | |

| FT | ACT SITE | 747 | 747 | BY SIMILARITY. |
|-----------------------|--|----------------------|--------------------|---|
| FT | ACT SITE | 1103 | 1103 | BY SIMILARITY. |
| SQ | SEQUENCE | 1458 AA; | 161343 MW; | 9D1608F47B1062B6 CRC64; |
| Query Match | | 77.1%; Score 5987.5; | DB 1; Length 1458; | |
| Best Local Similarity | | 77.6%; Pred. No. 0; | | |
| Matches | 1130; Conservative | 127; Mismatches | 197; Indels | 3; Gaps |
| Qy | 1 MGLRPGRIFLLELLMLLGOCPTQHNTSPRKISTLSQCOLWPELTQNSPPCNPKUNKGVNMPSK | 60 | Db | 1080 NSVPISVHQLRADIKVVAQALGSDLTAVGARNPSDLPSTSWAGLSNIGDGSNLBHT |
| Qy | 1 MALWPVSFVLGLPLIGRGAQDQIOTSSGKNTLEGQWLPESIKTFPPCIPDKTIAESVSE | 60 | Db | 1020 LEPGKKTDAIDLTAITLCTPONEPFLTRFRNSDVTYTPSPRAVENKNSDFLTAWN |
| Qy | 61 SVISLKPSDIFKVAIGNLNPPIPPIPGTGLEKDWTE RQOQCMGVMWTSIDRYSP | 119 | Db | 1140 TLPNLKKKNPSLIGPFTGIFTGEGMGLNAYVQJARAQDMPAQADLVERMKNSPDINLK |
| Qy | 61 SVHSRISLPSDIFKVAIGNVETAPDGGADLEBODTCRKEPQACMGVWVTLSDIGRFSP | 120 | Db | 1140 TLPNLKKKNPSLIGPFTGIFTGEGMGLNAYVQJARAQDMPAQADLVERMKNSPDINLK |
| Qy | 120 SVPMPVCHTGKRVIFHDGAEDLWTOQAELVNRMKENLQDFDWKLINVFSSNSAQCYL | 179 | Db | 1140 TLPNLKKKNPSLIGPFTGIFTGEGMGLNAYVQJARAQDMPAQADLVERMKNSPDINLK |
| Qy | 120 SVPMPVCHTGKRVIFHDGAEDLWTOQAELVNRMKENLQDFDWKLINVFSSNSAQCYL | 179 | Db | 1140 TLPNLKKKNPSLIGPFTGIFTGEGMGLNAYVQJARAQDMPAQADLVERMKNSPDINLK |
| Qy | 121 SALMPLC_PETRLVPRGGAEGLWQATELYRSKRBNPQDFERDHWKLINVFFSNTSQCP | 179 | Db | 1140 TLPNLKKKNPSLIGPFTGIFTGEGMGLNAYVQJARAQDMPAQADLVERMKNSPDINLK |
| Qy | 180 CPSAQONGLAAGGSVDELMCGVLDYQLOQEVRAPVNLDSVEAEVSRQVHTWLSPAPER | 239 | Db | 1200 DWKUVTLFPGNDLHYCENPAHLATEVQHQIQLDLSIEELPRAFYVNEVELASL |
| Qy | 180 CPSAQOKGLVLGKMDKLTRTLQDYLQEVPKAFVNVLVIDSELAAFSRWRGAQSLPAAEP | 239 | Db | 1200 DWKUVTLFPGNDLHYCENPAHLATEVQHQIQLDLSIEELPRAFYVNEVELASL |
| Qy | 240 NCEBTTRILAKUVQWSTQEAWSLASSRYSECFSTVQOPPFYETPSLSEDPRIQ | 299 | Db | 1200 DWKUVTLFPGNDLHYCENPAHLATEVQHQIQLDLSIEELPRAFYVNEVELASL |
| Qy | 240 RCLRETSOLTKVLTOWSYLEAWSLASSKYNTQESFAVTFQPFYEESSLALSALIAEPLQ | 299 | Db | 1260 YQOGOGKCA_MLAQNQNCIRAHQSSEERKLVNNWQHGSFSFWSWHQYOREFDA |
| Qy | 300 DSTLILAWHILWNRMPEPAGEDEPLSVKHPMRMKCPSQESPYLSYRNNSYLRIQPKDQ | 359 | Db | 1260 YQOGOGKCA_MLAQNQNCIRAHQSSEERKLVNNWQHGSFSFWSWHQYOREFDA |
| Qy | 300 DPTTLALSIWNRMPEPIGKKEPBKERRKPLRCPTQSBSPYLYTRNSQQLTRUSQOCK | 359 | Db | 1260 YQOGOGKCA_MLAQNQNCIRAHQSSEERKLVNNWQHGSFSFWSWHQYOREFDA |
| Qy | 360 LEVREGAETIRCPCDKPSDTWPTSVRKLKDADIVGAMGDSLTAGNGACSQGNILDWLT | 419 | Db | 1260 YQOGOGKCA_MLAQNQNCIRAHQSSEERKLVNNWQHGSFSFWSWHQYOREFDA |
| Qy | 360 LEVREGTEIRCPDKPSDVSPTSVRKLKDADIVGAMGDSLTAGNGACSQGNILDWLT | 419 | Db | 1260 YQOGOGKCA_MLAQNQNCIRAHQSSEERKLVNNWQHGSFSFWSWHQYOREFDA |
| Qy | 420 QYGLSWSVGGDENTGTWTANLIREPFLSLKGFSSVICKETSPPAFLNQAV/GGRAED | 479 | Db | 1319 WVOOPFFONTLTPLANEGRGTDLTFSEDCPFHFSRGRHAMAIALWNMILEPVGRKTTSN |
| Qy | 420 QYRGLSWSVGGDQNSTVTLANLIREFNPSCFGVSGFRETTSQAFFNOAVAGRADG | 479 | Db | 1319 WVOOPFFONTLTPLANEGRGTDLTFSEDCPFHFSRGRHAMAIALWNMILEPVGRKTTSN |
| Qy | 480 LPVQARRLVDIMKNDTRHFQEDWMIKTFIGNDLCDFCNDLYHYSPOFNFTDNGKALD | 539 | Db | 1320 VVVOOPFFONTLTPLANEGRGTDLTFSDDCPFHFSRGRHAMAIALWNMILEPVGRKTTSN |
| Qy | 480 LIOQARLVALMNDTRINQEDWMIKTFIGNDLCDFCNDPYRYSPOFNFTDNGKALD | 539 | Db | 1320 VVVOOPFFONTLTPLANEGRGTDLTFSDDCPFHFSRGRHAMAIALWNMILEPVGRKTTSN |
| Qy | 540 ILHAEIPRAFPVNVLVKVLEISKLRELYQETKVSICPRMILRSICPCVLUKFEDNSTBASLIE | 599 | Db | 1379 PTHSRAKUJCPSPSPSVPLTYLNRSLPQOAEEARPEVLYWAVPAVAGVGLVNGIGTVW |
| Qy | 600 FNIKKQEKHOLLESGRVTREDTYYVQPFENDPMPKTSSEGJLADNSFAPDCFFRSSK | 659 | Db | 1379 PTHSRAKUJCPSPSPSVPLTYLNRSLPQOAEEARPEVLYWAVPAVAGVGLVNGIGTVW |
| Qy | 600 TIKEYERTFOOLDSDGRTDRDFTTQLOPFFEKVNMKPTQDGDFDNSEFAPCFFRSSK | 659 | Db | 1380 FTYSTRKLKCPSPSPSVPLTYLNRSLPQOAEEARPEVLYWAVPAVAGVGLVNGIGTVW |
| Qy | 660 SHRARSALWNMILEPVGORTTRKUFKENKINITCPNQVOPFLRTYKNSMOGHGTWLPCKD | 719 | Db | 1439 CRRCRRGRDPPLMLT |
| Qy | 660 AHAAHASLWNMILEPVGOKTTRHDFEGAVNITCPNQVOPFLSTYKNSVOQFGFWLPCRD | 719 | Db | 1439 CRRCRRGRDPPLMLT |
| Qy | 720 RAPSALHPTSVHALRADIOVWAALGDSLTAGNGISKPDPDLPVTTORGLSYSAAGDG | 779 | Db | 1440 RGMCRCPREDPPLSLIST |
| Qy | 720 RSPSASPPSVHALRADIOVWAALGDSLTAGIGIGSKPKNDLSDGTTORGLSYSSGCGD | 779 | Db | 1440 RGMCRCPREDPPLSLIST |
| Qy | 780 SLENUTTLPNLIBFNRNLTGYANTGANDTMAFLNQAVPGAKAEDLMSQVQTMQKHK | 839 | Db | 1440 RGMCRCPREDPPLSLIST |
| Qy | 780 SLNUTTLPNLIBFNRNLTGYANTGANDTMAFLNQAVPGAKAEDLMSQVQTMQKHK | 839 | Db | 1440 RGMCRCPREDPPLSLIST |
| Qy | 840 DDHRVNFHBDWKVITVQIGASDLCYDCTSDNSLYSAFNFYDHMLADYHREPREVNL | 899 | Db | 1440 RGMCRCPREDPPLSLIST |
| Qy | 840 DDHRVNFHBDWKVITVQIGASDLCYDCTSDNSLYSAFNFYDHMLADYHREPREVNL | 899 | Db | 1440 RGMCRCPREDPPLSLIST |
| Qy | 900 VDFLNPTIMQVFLGPNPKCPVQDQASVQCNVULRENSOELARLEAFSRAYSSMRELV | 959 | Db | 1440 RGMCRCPREDPPLSLIST |
| Qy | 900 VDFLNPTIMQVFLGPNPKCPVQDQASVQCNVULRENSOELARLEAFSRAYSSMRELV | 959 | Db | 1440 RGMCRCPREDPPLSLIST |
| Qy | 960 GSGRYDQTQDFSVVLLQPFQFQNIQFLVFLADGLPDTSSFAFDCHIHNQKFSQLARALWTM | 1019 | Db | 1440 RGMCRCPREDPPLSLIST |
| RESULT 2 | | | | |
| YAMB_SCHPO | STANDARD; | PRT; | 3655 AA. | |
| ID | YAMB_SCHPO | STANDARD; | PRT; | 3655 AA. |
| AC | Q10064; | | | |
| DT | 01-FEB-1996 (Rel. 33, Created) | | | |
| DT | 01-FEB-1996 (Rel. 33, Last sequence update) | | | |
| DT | 28-FEB-2003 (Rel. 41, Last annotation update) | | | |
| DE | Hypothetical protein CIP5.11C | | | |
| GN | SPAC1F5_11C | | | |
| OS | Schizosaccharomyces pombe (Fission Yeast) | | | |
| OC | Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes; | | | |
| OC | Schizosaccharomycetales; Schizosaccharomycetaceae; | | | |
| OX | NCBI_TaxID=4896; | | | |
| RN | [1] | | | |
| SEQUENCE FROM N.A. | | | | |
| STRAIN=972; | | | | |
| RX | LINE=21849401; PubMed=11859360; | | | |
| RA | Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., | | | |
| RA | Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman D., | | | |
| RA | Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., | | | |
| RA | Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A., | | | |
| RA | Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgeson G., | | | |
| RA | Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hodgeson G., | | | |
| RA | Jones K., Jones L., Jones M., Leathem S., McDonald S., McLean J., | | | |
| RA | Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., | | | |
| RA | Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbowitze B., | | | |
| RA | Rutherford K., Rutte S., Saunders D., Seeger K., Sharp S., | | | |
| RA | Skelton J., Simmonds M., Squares R., Squares S., Stevens K., | | | |
| RA | Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., | | | |
| RA | Woodward J., Voilaert G., Aert R., Robben J., Grymonprez B., | | | |
| RA | Weitjens I., Vansreels E., Rieger M., Schaefer M., Mueller-Auer S., | | | |
| RA | Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilpert H., | | | |
| RA | Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M., | | | |
| RA | Eger P., Zimmermann W., Wedler H., Wambutt R., Purcell B., | | | |
| RA | Goffeau A., Cadieu E., Dreano S., Goux S., Lelaike V., Mottier S., | | | |
| RA | Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., | | | |

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Hong T.B., Scheetz T.B.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Brownstein M.J., Usdin T.B., Tashiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loqueilano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Kettenman M., Madan A., Rodriguez S., Sanchez A.,
 RA Fahey J., Helton B., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimmold J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smilus D.E.,
 RA Schmerch A., Schein J.B., Jones S.J.M., Marra M.A., Smailus D.E.,
 RA "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RL [5]
 RP SQUENCE OF 491-571 AND 610-671 FROM N.A.
 RX MEDLINE=98389634; PubMed=712139;
 RA Centra M., Memeo B., D'Apolito M., Savino M., Tanzano L.,
 RA Notarangelo A., Liu J., Doggett N.A., Zelante L., Savoia A.;
 RT "Fine exon-intron structure of the Fanconi anemia group A (FAA) gene
 and characterization of two genomic deletions";
 RL Genomics 51:461-467(1998).
 RN [6]
 RP SEQUENCE OF 491-542 FROM N.A.
 RX MEDLINE=98375690; PubMed=711872;
 RA Levran O., Doggett N.A., Auerbach A.D.;
 RT "Identification of Alu-mediated deletions in the Fanconi anemia gene
 RNA."; Hum. Mutat. 12:145-152(1998).
 RN [7]
 RP SUBCELLULAR LOCATION AND MUTAGENESIS.
 RX MEDLINE=98141611; PubMed=9742112;
 RA Naef D., Kupfer G.M., Suliman A., Lambert K., D'Andrea A.D.;
 RT "Functional activity of the Fanconi anemia protein FAA requires FAN
 binding and nuclear localization";
 RL Mol. Cell. Biol. 18:5952-5960(1998).
 RN [8]
 RP PHOSPHORYLATION.
 RX MEDLINE=99007271; PubMed=9789045;
 RA Yamashita T., Kupfer G.M., Naef D., Suliman A., Joenje H., Asano S.,
 D'Andrea A.D.;
 RA "The Fanconi anemia pathway requires FAA phosphorylation and FAA/FAC
 RT nuclear accumulation";
 RT Proc. Natl. Acad. Sci. U.S.A. 95:13085-13090(1998).
 RL [9]
 RP VARIANTS FA, AND VARIANTS.
 RX MEDLINE=98058767; PubMed=971789;
 RA Levrin O., Brilich T., Magdalena N., Gregory J.J., Batish S.D.,
 RA Verlander P.C., Auerbach A.D.,
 RT "Sequence variation in the Fanconi anemia gene FAA";
 RT Proc. Natl. Acad. Sci. U.S.A. 94:13051-13056(1997).
 RL [10]
 RP VARIANTS FA PRO-110 AND GLY-117.
 RX MEDLINE=9925041; PubMed=10210316;
 RA Kupfer G., Naef D., Garcia-Higuera I., Wasik J., Cheng A.,
 RA Yamashita T., Tipping A., Morgan N., Mathew C.G., D'Andrea A.D.;
 RT "A patient-derived mutant form of the Fanconi anemia protein, FANCA,
 RT is defective in nuclear accumulation";
 RT Exp. Hematol. 27:587-593(1999).
 -I- FUNCTION: DNA repair protein that may operate in a postreplication
 repair or a cell cycle checkpoint function. May be involved in
 interstrand DNA cross-link repair and in the maintenance of normal
 chromosome stability.
 CC -I- SUBUNIT: Belongs to a multisubunit complex composed of FANCA,
 CC FANCC, FANCF and FANCG proteins.
 CC -I- SUBCELLULAR LOCATION: Nuclear (major) and cytoplasmic (minor).
 CC -I- ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=2;
 Name=1; Iso1=015360-1; Sequenced=Displayed;
 CC Name=2;
 CC

CC I-B01D=015360-2; Sequence=VSP_007039;
 CC Note=No experimental confirmation available;
 CC -I- PRM: Phosphorylated primarily on serine residues. Phosphorylation
 is required for the formation of the nuclear complex. Not
 phosphorylated in cells derived from groups A, B, C, E, F, G, and
 H.
 CC -I- DISBASE: Defects in FANCA are the cause of one of the 8
 complementation groups of Fanconi anemia (FA) [MIM:227550], an
 autosomal recessive disorder characterized by progressive
 pancytopenia, a diverse assortment of congenital malformations,
 and a predisposition to the development of malignancies. At the
 cellular level it is associated with hypersensitivity to DNA-
 damaging agents, chromosomal instability (increased chromosome
 breakage), and defective DNA repair. FANCA defects account for 65%
 of FA.
 CC -I- DATABASE: NAME=Fanconi Anemia Mutation Database;
 CC WWW="http://www.infobiogen.fr/services/chromcancer/genes/FAID102.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.ebi-sib.ch/announce/
 CC or send an email to license@ebi-sib.ch).
 CC -----
 DR EMBL; X99236; CAB67610.1; --.
 DR EMBL; 283067; CAB0545.1; --.
 DR EMBL; 283068; CAB0545.1; JOINED.
 DR EMBL; 283059; CAB0545.1; JOINED.
 DR EMBL; 283070; CAB0545.1; JOINED.
 DR EMBL; 283071; CAB0545.1; JOINED.
 DR EMBL; 283072; CAB0545.1; JOINED.
 DR EMBL; 283073; CAB0545.1; JOINED.
 DR EMBL; 283074; CAB0545.1; JOINED.
 DR EMBL; 283075; CAB0545.1; JOINED.
 DR EMBL; 283076; CAB0545.1; JOINED.
 DR EMBL; 283077; CAB0545.1; JOINED.
 DR EMBL; 283078; CAB0545.1; JOINED.
 DR EMBL; 283079; CAB0545.1; JOINED.
 DR EMBL; 283080; CAB0545.1; JOINED.
 DR EMBL; 283081; CAB0545.1; JOINED.
 DR EMBL; 283082; CAB0545.1; JOINED.
 DR EMBL; 283083; CAB0545.1; JOINED.
 DR EMBL; 283084; CAB0545.1; JOINED.
 DR EMBL; 283085; CAB0545.1; JOINED.
 DR EMBL; 283086; CAB0545.1; JOINED.
 DR EMBL; 283087; CAB0545.1; JOINED.
 DR EMBL; 283088; CAB0545.1; JOINED.
 DR EMBL; 283089; CAB0545.1; JOINED.
 DR EMBL; 283090; CAB0545.1; JOINED.
 DR EMBL; 283091; CAB0545.1; JOINED.
 DR EMBL; 283092; CAB0545.1; JOINED.
 DR EMBL; 283093; CAB0545.1; JOINED.
 DR EMBL; 283094; CAB0545.1; JOINED.
 DR EMBL; 283095; CAB0545.1; JOINED.
 DR EMBL; 283096; CAB0545.1; JOINED.
 DR EMBL; 283097; CAB0545.1; JOINED.
 DR EMBL; 283098; CAB0545.1; JOINED.
 DR EMBL; 283099; CAB0545.1; JOINED.
 DR EMBL; 283100; CAB0545.1; JOINED.
 DR EMBL; 283101; CAB0545.1; JOINED.
 DR EMBL; 283102; CAB0545.1; JOINED.
 DR EMBL; 283103; CAB0545.1; JOINED.
 DR EMBL; 283104; CAB0545.1; JOINED.
 DR EMBL; 283105; CAB0545.1; JOINED.
 DR EMBL; 283106; CAB0545.1; JOINED.
 DR EMBL; 283107; CAB0545.1; JOINED.
 DR EMBL; 283108; CAB0545.1; JOINED.
 DR EMBL; 283109; CAB0545.1; JOINED.
 DR EMBL; 283110; CAB0545.1; JOINED.
 DR EMBL; 283111; CAB0545.1; JOINED.
 DR EMBL; 283112; CAB0545.1; JOINED.
 DR EMBL; 283113; CAB0545.1; JOINED.
 DR EMBL; 283114; CAB0545.1; JOINED.
 DR EMBL; 283115; CAB0545.1; JOINED.
 DR EMBL; 283116; CAB0545.1; JOINED.
 DR EMBL; 283117; CAB0545.1; JOINED.
 DR EMBL; 283118; CAB0545.1; JOINED.
 DR EMBL; 283119; CAB0545.1; JOINED.
 DR EMBL; 283120; CAB0545.1; JOINED.
 DR EMBL; 283121; CAB0545.1; JOINED.
 DR EMBL; 283122; CAB0545.1; JOINED.
 DR EMBL; 283123; CAB0545.1; JOINED.
 DR EMBL; 283124; CAB0545.1; JOINED.
 DR EMBL; 283125; CAB0545.1; JOINED.
 DR EMBL; 283126; CAB0545.1; JOINED.
 DR EMBL; 283127; CAB0545.1; JOINED.
 DR EMBL; 283128; CAB0545.1; JOINED.
 DR EMBL; 283129; CAB0545.1; JOINED.
 DR EMBL; 283130; CAB0545.1; JOINED.
 DR EMBL; 283131; CAB0545.1; JOINED.
 DR EMBL; 283132; CAB0545.1; JOINED.
 DR EMBL; 283133; CAB0545.1; JOINED.
 DR EMBL; 283134; CAB0545.1; JOINED.
 DR EMBL; 283135; CAB0545.1; JOINED.
 DR EMBL; 283136; CAB0545.1; JOINED.
 DR EMBL; 283137; CAB0545.1; JOINED.
 DR EMBL; 283138; CAB0545.1; JOINED.
 DR EMBL; 283139; CAB0545.1; JOINED.
 DR EMBL; 283140; CAB0545.1; JOINED.
 DR EMBL; 283141; CAB0545.1; JOINED.
 DR EMBL; 283142; CAB0545.1; JOINED.
 DR EMBL; 283143; CAB0545.1; JOINED.
 DR EMBL; 283144; CAB0545.1; JOINED.
 DR EMBL; 283145; CAB0545.1; JOINED.
 DR EMBL; 283146; CAB0545.1; JOINED.
 DR EMBL; 283147; CAB0545.1; JOINED.
 DR EMBL; 283148; CAB0545.1; JOINED.
 DR EMBL; 283149; CAB0545.1; JOINED.
 DR EMBL; 283150; CAB0545.1; JOINED.
 DR EMBL; 283151; CAB0545.1; JOINED.
 DR EMBL; AC005360; AAC28751.1; --.
 DR EMBL; AC005565; AAC33304.1; --.
 DR EMBL; AC005567; AAC33401.1; --.
 DR EMBL; BC005979; AAH08979.1; --.
 DR EMBL; AJ225084; CAAI2393.1; --.
 DR EMBL; AU225085; CAAI2394.1; --.
 DR EMBL; AF04569; AAC28331.1; --.
 DR Gene; HGNC:3582; FANCA.
 DR MM; 607139; --.
 DR MIM; 227650; --.
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0006334; C:nucleus; TAS.
 DR GO; GO:000281; P:DNA repair; TAS.
 DR GO; GO:0006461; P:protein complex assembly; TAS.

DR InterPro; IPR003151; Farconia.

DR PRINTS; PF03511; Farconia; 1.

DR PRINS; PR0826; FANCONIA GENE.

KW DNA repair; Nuclear protein; Phosphorylation; Disease mutation; Polymorphism; Alternative splicing.

Query Match 1.8%; Score 138; DB 1; Length 1455;

Best Local Similarity 19.3%; Pred. No. 0.93; DB Matches 225; Conservative 139; Mismatches 393; Indels 408; Gaps 58;

QY 143 IQQELBLVRNKKENIQLDFQPDWKULINVFRNSNAQCQCYLCPAQDNLAAAGGVDLSMG--- 198

Db 366 ISABELVSHLQEVLTEQ-EHWRWVLVSFISALVCC- PEAOQ--LLEDWVARLMQAFFE 420

Qy 199 -----VLDYLQEVPRRAFLVNLSEVAEVS-RQHGTWLSAPERNCSET 245

Db 421 SCQLDMSVTAFLVLRQALEGPAFLSYADWFKAESFGSTRGYHG- CSKKA 469

Qy 246 -----TRIAKVTMOWSISQEAWSNILLASRYSEOFSTVVFQOPPFYET--TPSLHSEDRLL 298

Db 470 LVLPLFTSELVP-----FESPRYLQWHILHPLVPSKRSLLTDYISAKTRL 518

Qy 299 QDSTTLAHLWN---RMMEPAGEKDPLS-----VTKH-GPRMKCPSQESYLV 341

Db 519 AD---LKSIENNGLYEDISSAGDITEPHISQALQDVKAIMVFEHTGMNPVTMEASFR 575

Qy 342 FSYRNNSVLTIRLOKPKDQKLEVRGEAEIRCPDKDPSDTPTVSRHL-----KPADI 391

Db 576 RPYVSHFLPALLTTRVLPKVUPSRVAFIESLKRADKIPPSLVSYSTCQACSAEKKPDE- 634

Qy 392 NIVGALGDSLTAGNGAGSTPGNVLIDVLTQYRGLSWSVGEGDENIGTVTIANLIREFNSSL 451

Db 635 -----AALGVRAEPNS-----AEEPIQGQTLAIGER----- 661

Qy 452 KGPSSVGTTKEPSN-----AFLNQ---AVAGGRAEDLPYQARRVLDWMKNDRHF 499

Db 662 -----ASMDPSQDVSQAVISRLRAVGNLGHNEDSSWEISK----- 713

Qy 500 QEDWKITLFIGGNDLCDFCNDLVHYSQNFTDNIKGALDLILHNEVPRAFVNLTVBIV 559

Db 714 REHMAVOLI-----LTSFCQONIMMAS-----SVAP----- 738

Qy 560 NIRELKYQEKVVCPMLRSLC---PCVLFKFDNSTELATLTFENKFKQEKTHQL--- 611

Db 739 -----PERQGPWAHLFVRTMCVRVLPAVL-----TRICQLQRHQGPSLSAPHVGLAA 786

Qy 612 --IESGRYDTRDFTVYVQPFEEVNDMKTSEGIP-----DNSFAPCFC- 654

Db 787 LAVHLG--ERSALPEV-----DVGPAPAGLCPVPLALPSLTCRTRDSF--CLK 835

Qy 655 -----HFSKSHSRAASALWNMMLPEVGQKTRHKKENKNINTCPNOVQPFILT 703

Db 836 RCTPAATSYSLCKFSSQSRTDCSCLSPCLK-----KQFLMPRLFSEARQPLSEE 886

Qy 704 YKNSMQSHGWTWLPCRDRAPS-----HPTSVHRLRADIQVWAAGPSLTAENGIGSKPD 759

Db 887 DVASLSMRPLHPSADMORALSLWHTHTFREVLKEEDVHL--TYQDNLHLEIOPBHD 944

Qy 760 DLPDVTTQ-----YRGISYSGAGGSILENV-TLPNPLREFURNLGYAVNTGQD 807

Db 945 AUSDTTERDFOHAWAHEHFLPESSAAGCCGDIQACITLVNALMDPFISSRY----D 999

Qy 808 ANDTNAFLNQAVPGAKA--EDLMSQVQTLMKQKDDHRVNFFEDWKVITVLIGGSIDLQDY 865

Db 1000 HSEN---SDAFLFGGTRGEDIISRLQEWADLE----LQD---LIVPIGHTP--- 1042

Qy 866 CTDSNLVSAANFV-DHGRNALDVL-----HREV--PRVLAUNLDPFLNPTMR 909

Db 1043 -----SOEHLFLPEIFRRRLQALTSGWSVAASLQRQRELLMKTRILRL----- 1085

Qy 910 QVFLGNPDKCPVQOQASVLCNCVLTRENSQELARLEFSRATRSSMELVSGRYPTQED 969

Db 1086 -----PSVCGS--SFQEQPITRACEPFLHYNSEMAMFCNSHGGAL 1129

Qy 970 FSVUJQPFONIOLPVLAQGLPD--TSFAPDCCHPNQKPHSOLARL-WTNMLRPLGSK 1026

Db 1130 --ITAHHFRGLINACIRSRRDPSLMVDFIACK--QTKCPLITSALVWMPSPSLBV-- 1180

Qy 1027 TETDLRAEMPITCPTONPPRLPRNSNYTIPKAIE-NWGSDFLCTHE----- 1075

Db 1181 -----LLCRWRRUCQSPSLPRE-----LOKLOQGRQFASDFLPEAASPAPNPD 1223

Qy 1076 WKASNVSPTVSHOILRPADTK-----VVALGDSITSTAVGARPNN 1114

Db 1224 WLSAAALHFAIQQVREENIRKQLKKUDCEREELLVFLPFFSLMGLLHSLS-----NS 1277

Qy 1115 SDUPTSW-----RGLSW 1127

Db 1278 TDLPKAFHVCAATCLECLEKRKISW 1302

RESULT 4

GLI4_XENLA
ID GLI4_XENLA STANDARD PRT: 1361 AA.

AC 091661;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 11-SEP-2003 (Rel. 42, Last annotation update)

DB Zinc finger protein GLI4 (Neural specific DNA binding protein xGLI4) (xGLI-4).

GN GLI4.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus.

OC NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE:97346726; PubMed=9203143;

RA Maxine J.C., Belfroid E.J., Pendeville H., Martial J.A., Pieler T.

RT "A role for Xenopus GLI-type inc finger proteins in the early embryonic patterning of mesoderm and neuroectoderm.";

RL Mech. Dev. 63:211-225(1997).

CC -- FUNCTION: HAS AN ESSENTIAL ROLE IN THE EARLY EMBRYONIC PATTERNING OF MESODERM AND NEUROECTODERM.

CC -- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -- SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U42462; AAH98467.1; -.

DR PIR; T3084; T3084.

DR RHEP; P08151; 2GL2.

DR InterPro; IPR007087; znf_C2H2.

DR PFAM; PF0096; zf_C2H2; 5.

DR SMART; SM0055; Znf_C2H2; 5.

DR PROSITE; PS0028; ZINC_FINGER_C2H2_1; 4.

DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 5.

DR Zinc-finger; Metal-binding; DNA-binding; Transcription regulation;

KW Nuclear protein; Repeat.

FT ZN_FING 289 314 C2H2-TYPE.

FT ZN_FING 322 349 C2H2-TYPE.

FT ZN_FING 355 379 C2H2-TYPE.

FT ZN_FING 385 410 C2H2-TYPE.

FT ZN_FING 416 441 C2H2-TYPE.

SO SEQENCE 1361 AA; 149554 MW; 03AC398AP4CC4DC CRC64;

Query Match Best Local Similarity 1.7%; Score 131.5; DB 1; Length 1361;

| Matches | 249; Conservative | 140; Mismatches | 498; Indels | 379; Gaps |
|---------|--|-----------------|-------------|-----------|
| 56 | NMPSKSVHS---LKPSDIK-----VAIGNLEIPPPGTCGLEKQDWTERPQOC----- | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 10 | NSPTLSMISARGLSPAAEVAEHLKERRGIYGLAPPNGPT-----TPTEYCHQMA 59 | | | |
| 104 | -----MGMTVLS-----DIRYFSSVPMVWCHGKRVHIDGAE 139 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 60 | FLASHSPYCBLVGOSAAGNTSHLHDYLTPMDVSRSSPRVTPLR-----SRKRAISIPLS 117 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 140 | DWIOQELVRNKMENIQDQFDWKUNINFFSNASQCYLCPSAQNGLAGGV----- | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 118 | DASILQTMRTSPSL-----RAYINNSRSSASAGYGYHGNSGAISSPAFSP 165 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 195 | -----ELMGVDIYLOQEVPRFAVNVLIDLSEVAEVRSRQYHG-----TWLSPAPEPCNCSEETT 246 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 166 | PIHKPCSLSAALSQ-ORSLSSSGHTPLLRSPTFASRQQGALTSAAPAPSNNSSAPDS 224 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 247 | RLAKUVMQWYQEAMNLL--ASSRYSEQSFTVVF-QP----- | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 225 | VLNKVSYSESAVSNTVNQVIIHRSKVKTEERADSVRPPQDPHLDIKEDLDKDECKOPE 284 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 283 | FFYETTPSLHSEDPRLDQSTTLAWHINRNMRMEPAGEKDE-----PLSVRH----- | 327 | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 285 | HIVETNCHWQCGSKERFDTQDOLVHINDH-H-GKEKEFVCRWQDCRSREOKPFKQAOYML 342 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 328 | -----GRPMKCPHQ-----PSPYLFSYRNSN-----YLTRL 353 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 343 | VVHMRHHTGKPKHCTEEGFKAYSRLLENLKTHLRSHTGEKPYVCDHEGCNKAFASNDR 402 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 354 | OKPODKLEVREGAEI-RCPDKDPSDTVPTSV-----HULKPAJINVAGLDS 400 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 403 | AKHQRTHSNESNEKPYVICKVPGCCKTRVYDTPSSPKRKHVKUTHGPEAHYTKKHNDITQK-PS 460 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 401 | LTAGNAGGSTPGNVLDVLTQYRGLSMSVVGDE----NIGTVTTLANTLREBTFNPLSKGSV 456 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 461 | LPKENG----DNEASAKLSCREHSDVSRSRQDQEBCQLQTRIKTEDNMHHQSSP----GGQS 512 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 457 | GTKGETSPNAFLNQAVG-----GRAEDLPVQARRLVLDMEN-----D 494 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 513 | SCSSEPSPYGNNTNIDSGVDVSLAMQGSLGDLFGLBETSPVUDSTVSWORSGPATPET 572 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 495 | TRIHQE-----DWKITLFIGGNDICDFCNLDLVHYSQPONTDNGKALDILHAE 544 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 573 | QRIHSAETGTAAEREIKDNERFLITYEPNATCQ-NTRLIPTISANGFDDVIGVPPSILIN- 629 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 545 | VPRAF----VNUITVTELVNURELYOEKKVYCPRMILRLSLCPVCKVLKFDDNSTELATLIEFN 601 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 630 | -PRAELEMNDVUTMMQNLNERR-----DISTSL----- 657 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 602 | KKFQEKTQHOLIESGRYDRTEDFTWVQPFENVMKPTSE--GLPDNSFFAPDCFHSSK 659 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 658 | -----SSAYTSRR---SSGISPYFSSRRSSETSFQFGRLNNSSADSYDPISTD 703 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 660 | SHSRAASALMNMLERVGOKTRHKKEENKNITC-----PNQVQPLRTKQSMQ 709 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 704 | ASRRSEASQHSGLPMLNLUPTAQHYRLKAKVAATGCPPPTPLPNMDRIGLRNKLSIMD 763 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 710 | GHGTWLP-----CRDRAPS-LHPTSWHALRFADTOVVAALGDSLTAITAGNGIGSKP 758 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 764 | GADFPFLPPFROLVPPRCSDGGNAGLTPMWHF-PGNNSSRAS-----DPVRTAGDDK 820 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 759 | DDLPPDTVQYRGLSYAGGDSLENTL--PNLREFNRNLTGYAVGTGDAN----- 809 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 821 | --LPRFRSRFH-----SMNSNTLTHPPSLSRBRNGGLOHYTCSD3GLHRAVYSPR 867 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 810 | -----DTNAFL---NOAVPGAKAEDM---SQVQTMQKQKDHRWFHDKWVITLVI 857 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 868 | PPSISENVAMEAISCDADVPGCD-DLMLPDPDVQYRSQRNREAPCQNQTEYS----- 920 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 858 | GGSDLCDYCTSUNLYRA----NFVYHRLNADLVHREREVY-----LYNLFNLPNT 906 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |
| 921 | -----SPARLNQNTKSFHNTTPEQPRAGAYISRNFPALAECIGQTAGMQDMNPV 972 | | | |
| b | : : : : : : : : : : : : : : : : : | | | |

RESULT 5
 C2TA_MOUSE
 C2TA_MOUSE STANDARD; PRT; 1155 AA.
 ID P79621; 04787; 078036; 078109; Q21115; Q9TP1;
 AC 15-DEC-1998 (Rel. 37, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE MHC CLASS II transactivator (CITA).
 GN MHC2TA OR CITA OR C2TA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
 OX NCBI_TAXID=10090;
 RN [1]
 RP SOURCE FROM N.A. (ISOFORMS 1, 2 AND 3), TISSUE SPECIFICITY, AND
 RP INDUCTION.
 RC STRAIN=BALB/c;
 RX MEDLINE=97327562; PubMed=9184229;
 RA Muhlethaler-Mottet A., Otten L.A., Steinlein V., Mach B.;
 RT "Expression of MHC class II molecules in different cellular and
 functional compartments is controlled by differential usage of
 multiple promoters of the transactivator CITA";
 RL EMBJ. 16:2851-2860(1997).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=NOD; TISSUE=Spleen;
 RX MEDLINE=97152412; PubMed=8995190;
 RA Sims T.N., Elliott J.F., Ramasar V., Denney D.W. Jr., Halloran P.;
 RT "Mouse class II transactivator: cDNA sequence and amino acid
 comparison with the human class II transactivator.";
 RL Immunogenetics 45:220-222(1997).
 RN [3]
 RP SEQUENCE OF 955-1097 FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Thymus;
 RX MEDLINE=96188865; PubMed=8620527;
 RA Paneli M.C., Wang E., Shen S., Schluter S.P., Bernstein R.M.,
 Harsh E.M., Stoeck A., Gangavalli R., Barber J., Jolly D.,
 R.A. Aszpriore E.T.;
 RT "Interferon gamma (IFNgamma) gene transfer of an EMT6 tumor that
 poorly responsive to IFNgamma stimulation: increase in tumor
 immunogenicity is accompanied by induction of a mouse class II
 transactivator and class II MHC.";
 RL Cancer Immunol. Immunother. 42:99-107(1996).
 RT -1- FUNCTION: ESSENTIAL FOR TRANSCRIPTIONAL ACTIVITY OF THE HLA C
 II PROMOTER; ACTIVATION IS VIA THE PROXIMAL PROMOTER. NO DNA
 BINDING OF IN VITRO TRANSLATED CITA WAS DETECTED. MAY ACT IN
 COACTIVATOR-LIKE FASHION THROUGH PROTEIN-PROTEIN INTERACTIONS
 CONTACTING FACTORS BINDING TO THE PROXIMAL MHC CLASS II PROMO-
 TO ELEMENTS OF THE TRANSCRIPTION MACHINERY, OR BOTH. ALTERNAT-
 IT MAY ACTIVATE HLA CLASS II TRANSCRIPTION BY MODIFYING PROTE-
 THAT BIND TO THE MHC CLASS II PROMOTER (BY SIMILARITY).
 CC CC -1- ALTERNATIVE PRODUCTS:
 CC CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC CC Event=Alternative splicing; Named isoforms=3;
 CC Name1; Synonyms1;

CC ISO-1d-P79621-1; Sequence=Displayed;
 CC Name=2; Synonyms=III;
 CC IsoId=P79621-2; Sequence=VSP_007214, VSP_007215;
 CC Name=3; Synonyms=IV;
 CC IsoId-P79621-3; Sequence=VSP_007216;
 CC -I- TISSUE SPECIFICITY: Isoform 1 is expressed at very high levels in dendritic cells, at very low levels in spleen and thymus and is not detected in other tissues. Isoform 2 is detected at high levels in spleen and tonsil as well as in a number of B-lymphocyte cell lines, and at very low levels in dendritic cells.
 CC -I- INDUCTION: By interferon-gamma.
 CC -I- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
 CC -I- SIMILARITY: Contains 1 NACHT domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch.)

DR EMBL; AF000006; RAB92364.2; .
 DR EMBL; AF000007; RAB92365.1; .
 DR EMBL; AF042158; AAC34366.1; .
 DR EMBL; AF042159; AAC34367.1; .
 DR EMBL; AF107010; AAF0839.1; .
 DR EMBL; U60653; AAB4885.1; .
 DR EMBL; U46562; AAB0504.1; .
 DR MGI; MGI:108445; C2ta.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR PFam; PF00560; LRR; 1.
 DR SMART; SM0068; LRR; RI; 4.
 DR PROSITE; PS50837; NACHT; 1.
 KW Leucine-rich repeat; Repeat; Alternative splicing;
 FT REPEAT . 175 198
 FT 243 269
 FT DOMAIN . 439 749
 FT REPEAT . 979 1000
 FT REPEAT . 1008 1031
 FT REPEAT . 1039 1062
 FT DOMAIN . 129 215
 FT NP_BIND 445 452
 FT VARSPlic 1 77
 FT VARSPlic 78 94
 FT VARSPLIC 1 101
 FT CONFLICT 994 994
 SQ SEQUENCE 1155 AA; 127528 MW; F3FF05DEBBFB0C871 CRC64;

Query Match Best Local Similarity 22.9%; Prod. No. 3;
 Matches 102; Conservative 47; Mismatches 153; Indels 144; Gaps 26;

QY 39 ETIKKNSPPPCNPN-KLGVNMP-----SKSVHSLKPSDKFVVAIGNLEIPDPD----- 85
 Db 337 EFDITSPSPCGCPESSTKLKPWPAVERPQISQD--KYKA---LPSRSPRGFLVA 387
 QY 86 -----GTC-----DLEKQDWTERPQQVCMG---VMVTLSDITIRYFERSPVPHVC 126
 Db 388 VELYRARLERGSNSKQERELATPDWTER---QLAHGGLAEVQVQVSDCCR--PGETQVVA 442
 Qy 127 HIGK-----RVIPIHDGAELWIQAOBLVRNMKENQLQDFDWKLINTFP----- 171
 Db 443 VLGKAGOCKSHMARTVSH----TWACQO-----LQYDF-----WYYPVCHL 481

RESULT 6

YFAS_ECOLI ID YFAS_ECOLI STANDARD PRT: 1534 AA.
 AC P76464, P76465;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein yfas precursor.
 GN YFAS OR B2227/B2228.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
 OK NEBI_TaxID=562;
 RN [1]
 RP STRAIN_K12;
 RC STRAIN_K12 / MG1655;
 RK MEDLINE=9743617; PubMed=9778503;
 RK Blattner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mai B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN_K12;
 RK MEDLINE=9743617; PubMed=9205837;
 RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Mikki T., Mitsuhashi N.,
 RA Mizobuchi H., Mori H., Nakata S., Nakamura Y., Nashimoto H.,
 RA Ohshima T., Oyama S., Saito N., Sampei G., Satoh Y., Siva Sundaram S.,
 RA Tagami H., Takahashi H., Takeda J., Takekoto K., Uehara K., Wada C.,
 RA Yamagata S., Horikoshi T.;
 RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
 RT - K12 genome corresponding to 50.0-68.8 min on the linkage map and
 analysis of its sequence features.";
 RL DNA Res. 4:91-113(1997).
 RN [3]
 RP CONCEPTUAL TRANSLATION.
 RA Couderc E.;
 RL Unpublished observations (MAR-2002).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@ib-sib.ch).

DR EMBL; AE00312; AAC75288.1; ALT INIT.

DR EMBL; D90854; ; NOT ANNOTATED_CDS.

DR Ecogene; EGJ4081; yfas.

KW HYPOTHETICAL protein; Signal; complete proteome.

FT SIGNAL 1 HYPOTHETICAL PROTEIN YFAS.

FT CHAIN 39 1534 HYPOOTHEICAL PROTEIN YFAS.

SQ SEQUENCE 1534 AA; 169845 MW; 73B71431C51F8BC6 CRC64;

Query Match Best Local Similarity 1.6%; Score 126.5; DB 1; Length 1534; Matches 278; Conservative 18.8%; Pred. No. 5.9; Mismatches 556; Indels 473; Gaps 68;

QY 161 QFDWKLLINVFPSNASQCYLCPSAQONGA-----AGGVDELMGVLDYIQQE 206

Db 9 QFWHHL-SPKFGSGIAACISLSLVLGTLANADSLPSSVYAPPAGTTRFLADSSFSSE 67

QY 207 VPRAFVNVLVDSEAEVSRVYH-----GTMWSPAPRNCSEBETTLAKVUNOWSY-- 257

Db 68 EAK-----VRLAEPGRDVRYYQMEYGGDVDRVLYRIPDOPMAMFLRQOKNHLRIVQPOYIG 122

Qy 258 -----QEAWNSLASSRYSBQSFSPTVVFQPFYETPSLHSBDPRIODSTLAWHWN 310

Db 123 DGLANTLTWLWDNWYGSKRRVMORTFSSORNTQALPELQGNATIKPSR----YVON 178

QY 311 RMMPEPAGEKDPBLPSVKHGRPMKCPSQESPYLFSYRNNSNLRLQKPDQLEVERGAETRC 370

Db 179 NQPSPL--KKYPLVKQFRYPL---WQAKPF-----EPQGCVK-EGA--- 214

QY 371 PDKDPSDTVPTSVHRLKPADINVIGALGDSLTAENGAGSTPGNVLDVLTQYRGLSWSVGG 430

Db 215 -----SSNPISPQPQGNVYIPLQG----- 232

Qy 431 DENIGTVTTLNLTREFNPSLKGSFVGTCIKETSPNAFLNQAVAGGRADLPVQARRUDL 490

Db 233 -----QEPGLYLVEMYCG-----YRATTW 253

QY 491 MNQDTRIHFQEDWKITLFIGNGLCD-----FCNDLHVYSPQNFTDNIGKDALILHA 543

Db 254 FVSDTVALSKVSGKELLYWMTAGKKQGEARPGSELTWTDLGWMTRGVTDG-TLQLOHQ 312

QY 544 EVPRAFY-----NLVTLEIENLRELVEOKKVYCPRMILRSICPCVLUKFDDNSTELAL 597

Db 313 SPERSYILGKDAAEGGVFVSE----NFFYESEINTNRLYIFTDPRYRAGDRVVKV--- 364

Qy 598 IEFKKFQEKTHQ--LIESGRVDTREDFTVWVQPFENWDMPTKSEG--LDNSFFAP 651

Db 365 -TGREFHDPLHSSPIVSAPAKUSVLDANGSLIQTQVNTLDRANGQGSFRPENAVAG 422

QY 652 DCFHFSKSHSRAASALMNMLEP--VGQKTRHKFE-----NKINITCPN----- 695

Db 423 YEFLAYRQYQVYSSFRVANYIKHFEGLALAKKEPKYGEAVMSGKQIOLYDQEPVNA 482

QY 696 QVOPFLRTYKNSMOCGHGTWLPCDRDRAPSALHPTSVHALRPADTQVVAALKGDSLTAGNGTG 755

Db 483 RVQLSLRAQQLSMVGN----DLRYAGRFPVS-----LEGSETVSDASG 521

Qy 756 SKPDDLPVTTQYRGLSYISAGGDSLENVTPLMLREPNRNTGYAVT---GDANDT 811

Db 522 HVALNLPAPADKPSRVLTVSASDGAARYVTTKELIE--RGLAHYSLSTAQVNSNGS 579

QY 812 NAFLNAQAVFGAK-----AEDLMSQVQTMQMKMDHRVNF-----HEDW 850

Db 580 WVRYYAALBSSKQPVUTVWLRLDRTSISGEGLSSGGK-SFTVNEAKGCVNUTLDRDG 638

QY 851 KVITVL--IGG-----SDLCDYCTDSNLYSAANFVDHRLNALDVHREVPVRLVNUD 901

Db 639 LILAGLSHAVSGKGKSTAHTGTVDIVADKTLQPG-----ETAKMLITPFB 683

Qy 902 FLNPTIMROVFLGNDKCPVQOASVLCNCV--LTLR--ENSQELARLEAFRAYRSSMRE 957

Db 684 PIDBALL-----TLERDRVEQOSLISHSPANWMLTLQRINDTOEARVEV----- 726

QY 958 LVGSGRVTQEDFSVUL---QPFQVNLQPLVADGPDTSFAPPDCIHPNQKFHSQJAR 1013

Db 727 --SNSEPNITTSVLYRNGQTSFQNGQVAVPOL-----DRYVKTOKTHYQGE 775

QY 1014 ALWNTNMLPLGSKTETUDR-----AEMPTCPTQNEPLRTPRNSYT----- 1057

Db 776 LVNVELTSLIKGRPSAQLTGVDEMIALQPEIAPNIGKRFYPLGNNVRVTSSLIFI 835

QY 1058 -----YPIKAENNGSDFLCTEWKASNSNPTSVHOL-RPADIKY-VAAIGDSLWTA 1107

Db 836 SYDOALSSEBPVADG-----PAINRSERRVVKMLERPRREEVDTAAMPSLTD 881

QY 1108 VGARPNNNSDLP---TSRGLSNSIGCG-----NLETHITLPNLIKKFPRY 1151

Db 882 KOQKAYFTFLMPUSLTRWIRITARGWNGVGLGOGRAVYRSEKNLYMKMSMPY----- 935

Db 1152 LLGFSTSTWEGTAGLNVAEGARADMRAQAWDLVERMKNSPDINLKDWKLVTLFIGVN 1211

QY 936 -----RUGKPAAGLIFISQODNEP-----VALVTKFAG-- 964

QY 1212 DLCHYCENCEPAHLATEYV---OHIQOADDILSELBELPR-----AFV-NVNE/ME 1255

Db 965 --AEMRQQTITLHKGANYSILTONIQS-GLLSAELOQNGQVQDSISIYGLSFVDNSWPEVQ 1021

Db 1255 LASIYQOGCGKCMALQAQNCTCLRHSSOLEKELKVN-----W---NLQHGS 1303

QY 1022 QKNNMIGGCDNALMLPEQASNTRL--QSETQBIFBNNDLAVDEBPGVINTGSLI 1078

QY 1304 SFSY-W-----HOYTORBDPFAVWQPFQNTLPLNERG-DTDLTFPSDCFHSDRQHA 1356

Db 1079 PLSLAWRSLLADHQSAANDIRQMIQ--DNRALMQLGGPGRFTWIGED-----GNG 1128

QY 1357 EMALAWNNLPERVGRKITSNNFTHSRAKIKCPSPESTYLYTNRNSRLIPDQEEAP-- 1413

Db 1129 DAFLTAW-----AWYADMOASQAGIVTOQE--YWQHMLDS--YAEQADNMPMLH 1174

QY 1414 --EVLYW-----VPUVAAGVGLVYUIGVWMCRRGGRED 1448

Db 1175 RALVLTAWQEMNIPCKT--LILGDELAI--ARRGTCTED 1209

RESULT 7

SNE2_HUMAN ID SNE2_HUMAN STANDARD PRT; 6885 AA. Q8WWW5; Q8WXH1; AC Q8WXXH0; Q8NFS3; Q8NPF49; Q8TER7; Q8WWW3; Q8WWWW4; Q8WWWW5; Q8WWXH1; Q8WNU50; Q8UPO4; Q9Y2L4; Q9Y4R1; DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DE Nbsprin 2 (Nuclear envelope spectrin repeat protein 2) (Syne-2) (Synaptic nuclear envelope protein 2) (Nucleus and actin connecting DE element protein) (NUANCE protein).

GN SNE2 OR NUA OR KIAA1011.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI_TaxID=9606;

[1] SEQUENCE FROM N.A. (ISOFORMS 1 AND 9), FUNCTION, CHARACTERIZATION, AND INTERACTION WITH F-ACTIN.

RE MEDLINE-22113122; PubMed=12118075;

RA Zhen Y.-Y., Libotte T., Munck A.A., Korenbaum E.; RT "NUANCE, a giant protein connecting the nucleus and actin cytoskeleton"; J. Cell Sci. 115:3207-3222 (2002).

RA [2] SEQUENCE FROM N.A. (ISOFORMS 4, 5 AND 7).

RE MEDLINE-21652858; PubMed=11792844.

RA Zhang Q., Skoppe J.N., Yang F., Davies J.D., Hegyi L., Roberts R.G., Weissberg P.L., Ellis J.A., Shanahan C.M.,

- RT Nesprins: a novel family of spectrin-repeat-containing proteins that
RT localize to the nuclear membrane in multiple tissues.";
- RL *J. Cell Sci.* 114:4485-4498(2001).
- RN [3]
- RP SEQUENCE FROM N.A. (ISOFORM 1);
- RX MEDLINE=22286983; PubMed=12408964;
- RA Zhang Q., Ragnauth C., Greener M.J., Shannah C.M., Roberts R.G.,
RT "The nesprins are giant actin-binding proteins, orthologous to
RT Drosophila melanogaster muscle protein MSP-300.";
- RL *Genomics* 80:473-481(2002).
- RN [4]
- RP SEQUENCE FROM N.A. (ISOFORM 6).
- RC TISSUE=Tebis;
- RA Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
- RN [5]
- RP SEQUENCE FROM N.A.
- RX PUBMED=1258121;
- RA Heilig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C.,
RA Cattolico L., Levy M., Madan A., Mairey-Estrada B., Mangonot S.,
RA Pellecier E., Vico V., Anthouard V., Rowen L., Madan A., Oin S.,
RA Sun H., Du H., Pepin K., Artiguenave F., Robert C., Cruaud C.,
RA Brueles T., Jaillon O., Briez-Silla S., Combette S.,
RA Cure S., Segurens B., Aniere F., Samain S., Crespeau H., Abbassi N.,
RA Alach N., Boscos D., Dickhoff R., Dors M., Dubois I., Friedman C.,
RA Gouvenoux M., James R., Madan A., Mairey-Estrada B., Mangonot S.,
RA Martins N., Menard M., Ozras S., Ratcliff B., Trask B.,
RA Vacherie B., Bellmere C., Belser C., Bernard-Gomot M.,
RA Bartoli-Mavel D., Bouard M., Briez-Silla S., Combette S.,
RA Dutfosse-Lavrent V., Farron C., Lechaplais C., Louesse C., Muelalet D.,
RA Magdeleinat G., Pateau E., Petit E., Sirvain-Trukniewicz P., Trybou A.,
RA Vega-Czarny N., Batallie E., Bluet B., Bordelais I., Dubois M.,
RA Dumont C., Guerin T., Haffray S., Hammadi R., Muanda J., Pelouin V.,
RA Robert D., Wunderle E., Gauguet G., Roy A., Sainte-Marthe L.,
RA Verdier J., Verdier-Discalca C., Hillier L., Fulton L., McPherson J.,
RA Matsuda F., Wilson R., Scarpelli G., Gyabay G., Wicker P., Saurin W.,
RA Quetier F., Waterston R., Hood L., Weissenbach J.;
RT "The DNA sequence and analysis of human chromosome 14.";
RL *Nature* 421:601-607(2003).
- RN [6]
- RP SEQUENCE FROM N.A. (ISOFORM 8).
- RC TISSUE=Brain;
- RX MEDLINE=22288257; PubMed=12477332;
- RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Halek P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Seares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein B.M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loqueland N.A., Peers D., Abramson R.D., Mullah S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalobos D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boutefaud G.G.,
RA Blakesley R.W., Toucman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Snailius D.E.,
RA Schnecker A., Schein J.-E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.;"
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
- RN [7]
- RP SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 1-956 AND 5133-6895
- RC TISSUE=Spleen, and Tongue;
- RA Kikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.,
RA Niimotsu K., Wagatsuna K., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamibara K.,
RA Katsuma N., Sato K., Tanikawa M., Yanazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isomoto Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA FROM N.A.
- RC
- RA Matsuo K., Nakamura Y., Sakine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takehashi-Fujii A., Osima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahashi K., Matsuo Y., Nagai K., Isogai T.,
RT "NEO human cDNA sequencing project.";
- RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
- RN [8]
- RP SEQUENCE OF 5709-6895 FROM N.A. (ISOFORM 2).
- RC TISSUE=Brain;
- RX MEDLINE=99246063; PubMed=10231032;
- RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hiroshima M.,
RA Miyajima N., Tanaka A., Kocani H., Nomura N., Ohara O.,
RT "Prediction of the coding sequences of unidentified human genes. XIII.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro.";
- RL Res. 6:63-70(1999).
- RN [9]
- RP REVISIONS.
- RC TISSUE=Brain;
- RX MEDLINE=22158633; PubMed=12168954;
- RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.,
RT "Construction of expression-ready cDNA clones for KIAA genes: manual
curation of 330 KIAA cDNA clones.";
- RL DNA Res. 9:99-106(2002).
- RN [10]
- RP SEQUENCE OF 5754-6895 FROM N.A.
- RC MEDLINE=21154917; PubMed=11230166;
- RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ausurge W., Boehler M., Blocker H., Bauersachs S., Blum H.,
RA lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.-W., Ottewangler B., Obermaier B., Tempé J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RL analysis of 500 novel complete protein coding human cDNAs.";
CC Genome Res. 11:422-435(2001).
- I- FUNCTION: Involved in the maintenance of nuclear organization and
CC structural integrity. Probable anchoring protein which tethers the
CC nucleus to the cytoskeleton. Connects nuclei to the cytoskeleton
CC by interacting with the nuclear envelope and with F-actin in the
CC cytoplasm.
- I- SUBUNIT: Interacts with F-actin via its N-terminal domain.
CC -I- SUBCELLULAR LOCATION: Type IV membrane protein (potential). The
CC largest part of the protein is cytosolic, while its C-terminal
CC part is associated with the nuclear envelope, most probably the
CC outer nuclear membrane. Remains associated with the nuclear
CC envelope during its breakdown in mitotic cells.
- I- ALTERNATIVE PRODUCTS:
- CC Name=1; IsoID=Q8WXH0-1; Sequence=Displayed;
- CC Name=2; IsoID=Q8WXH0-2; Sequence=VSP_007164, VSP_007166;
- CC Note=No experimental confirmation available;
- CC Name=3; Event=Alternative splicing; Named isoforms=9;
- CC IsoID=Q8WXH0-3; Sequence=VSP_007155;
CC Note=Produced by exon skipping that results in a framehift. No
CC experimental confirmation available;
- CC Name=4; Synonyms=Beta;
CC IsoID=Q8WXH0-4; Sequence=VSP_007156;
- CC Name=5; Synonyms=Alpha;
- CC IsoID=Q8WXH0-5; Sequence=VSP_007157, VSP_007164, VSP_007165;
CC Name=6; IsoID=Q8WXH0-6; Sequence=VSP_007158, VSP_007165, VSP_007166;
- CC Note=No experimental confirmation available;
- CC Name=7; Synonyms=Gamma;
CC IsoID=Q8WXH0-7; Sequence=VSP_007154, VSP_007163;
- CC Name=8; IsoID=Q8WXH0-8; Sequence=VSP_007161, VSP_007162;
- CC Note=No experimental confirmation available;
- CC Name=9; Synonyms=NUANCE-N3;
- CC IsoID=Q8WXH0-9; Sequence=VSP_007159, VSP_007160;
- CC -I- TISSUE SPECIFICITY: Widely expressed, with higher level in kidney,
adult and fetal liver, stomach and placenta. Isoform 5 is highly expressed in
skelatal muscle and brain. Isoform 5 is highly expressed in

CC - I- pancreas, skeletal muscle and heart.
 CC - I- DOMAIN: The Klarischt domain mediates the nuclear envelope targeting.
 CC - I- SIMILARITY: Belongs to the Nesprin family.
 CC - I- SIMILARITY: Contains 1 actin-binding domain, (CH) domains.
 CC - I- SIMILARITY: Contains 1 Klarischt domain.
 CC - I- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF435010; AAL33547.1; --.
 DR EMBL; AF435011; AAL33548.1; --.
 DR EMBL; AY061757; AAL33800.1; --.
 DR EMBL; AY061758; AAL33801.1; --.
 DR EMBL; AY061759; AAL33802.1; --.
 DR EMBL; AF49911; AAN6043.1; --.
 DR EMBL; AY117404; CAB5505.1; --.
 DR EMBL; AL162032; ; NOT_ANNOTATED_CDS.
 DR EMBL; AL355094; ; NOT_ANNOTATED_CDS.

Query Match 1.6%; Score 126.5; DB 1; Length 6885;
 Best Local Similarity 17.6%; Prc. No. 67; Mismatches 408; Indels 431; Gaps 56;
 Matches 219; Conservative 174; Mimatches 408; Indels 431; Gaps 56;

Qy 342 FSYRNSNLYLTROKPODKLEVRREGAEIICPDOKPSDTPV----TSVRLKPADINVGA 396
 Db 5598 FLYCCEKNWOLKEEALKV-----DVANSPLPELUQSKQTYKMLAEVSINT 5646

Qy 397 LGISLTLTAGNGAGSTPGNVLDV-----LTOYRGLS--WS---VGGDENIGTVTLAN 442
 Db 5647 LADSYV-----TOSLQDPTTEIENRPFETTSKLFDRWAVQAVGTYQRQKGDVGVR 5700

Qy 443 ILREBFNPNSLKGSFGVTGKETSNAFLNQAVAGGERAEDLPVQASRLVLDIMKDNTRIHF-- 499
 Db 5701 QWQDFITSVENLP---RPLTDTHLSUVAKGQERFSL-YQTRSLIHILKN KEIHFORR 5754

Qy 500 -----
 Db 5755 RTTCALTLEAGEKULLTDLKTKESVGRRIQLQDSWDKMEPOLAEMIKQFQSTVETWDQ 518

Qy 519 C-----NDL-----VHSPQNFTDNTSKALDIHA 543
 Db 5815 CEKKIKELKSRLQVYLAQSEDPPLPELHDLNHEKELIKEPLEOSLASWNTNL-KELOTNA 5873

Qy 544 EVIRAFVNLUVTIVIENIRE---LYQERKVYCPMILRSCLPCVLFKDDNSTELATLIE 599
 Db 5874 DLTRH---VLVEDVMVJKEQTEHILRQMDCLRVRAIK---QEIDRINTVV 5921

Qy 600 FNKFKFOE-----KTHOLISGSGRYDFTWVQOPFEN-VDMPTKS 640
 Db 5922 FNEKKNKELCAWLQYOMENKVQLOTAISIBEMEKLQKOCMEETNU---FSENKLQKONG 5977

Qy 641 EGIPDONSFAPDCFHFSKSHSRA-----SALMNMLPEPVGQKTRHKPENKINI 691
 Db 5978 DQI-----IKASNKRSAEIDDKLNKINDRQHDFVIGSSVKKL----- 6018

Qy 692 TCPNQVQPFRLRTYKNSMOCHGTMPCDRDAPSALHPTSYHALPADIQVAALGDSLTA 751
 Db 6019 ---ETFAFIQOLDKDNMNLRTMIA---RIESLS---KE---WVIVCDDOBIO 6060

Qy 752 NSIGSKPDPDLPYTOYRGSLSYAGGDSLENTLPIRNFNLITGAYGTGDAWT 811
 Db 6061 KRLAQOQLQRDIE-----QHSAG---VESVNPNCIDL-----LHDS 6094

Qy 812 NAFLNQA-----VPGAKAEDLMSQVQT-LMQMKDHDHRVNPHEDWK 851

RESULT 8

| ID | ATCL_YARLI | STANDARD: | PRT: | 928 AA. |
|----|--|-----------|------|---------|
| AC | 043108; | | | |
| DT | 15-DEC-1998 (Rel. 37, Created) | | | |
| DT | 15-DEC-1998 (Rel. 37, Last sequence update) | | | |
| DS | 15-SEP-2003 (Rel. 42, Last annotation update) | | | |
| GN | Calcium-translocating ATPase 1 (EC 3.6.3.8) (P-type calcium ATPase). PMT1 OR SCAL. | | | |
| OS | Yarrowia lipolytica (Candida lipolytica). | | | |
| OC | Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; | | | |
| OC | Saccharomyces; Dipodascaceae; Yarrowia. | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N. A. | | | |
| RK | MLINE=98121320; Pubmed=9461422; | | | |
| RA | Park C.S., Kim J.-Y., Crispino C., Chang C.C., Ryu D.D.Y.; | | | |
| RR | "Molecular cloning of Yipm1, a S. cerevisiae PMr1 homologue encoding a novel P-type secretory pathway Ca2+-ATPase, in the yeast Yarrowia lipolytica."; | | | |
| RT | Gene 206:107-116(1998). | | | |
| RC | -I- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS OF AT COUPLED WITH THE TRANSPORT OF CALCIUM. HAS A ROLE IN THE SECRETORY PATHWAY. | | | |
| CC | -I- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+)(Cis) = ADP + phosphate + Ca(2+)(Trans). | | | |
| CC | -I- SUBCELLULAR LOCATION: Integral membrane protein. Golgi. | | | |
| CC | -I- SIMILARITY: Belongs to the cation transport ATPases family (P-type ATPases). | | | |

```

Db          375 DLIANAVILKPGQSHTEADEKAVALK--NSVSLAMMLKVNLNCNSKFRNREAGHLVGNA 431
Qy          615 -----GTYPTREDFTVHQ-PFFENVKPRTSCLPDNS-----PFA 650
Db          432 TDIALIEVLDYFGLEDTRETRKRVAEVPEFSSRKRMLTSTTGDSSTPPMISVKGASEVIA 491
Oy          651 PDEPHFPPSSKSHSRQ-----ASALWNMLBPGQTRRKPFENKINTCPNOV 697
Db          492 PFCYIYCKDGKTRAPPNDMRKKVTEIASMSNDQRLINFAKYQKQYEE---GSEE 545
Oy          698 QPFRLTYKNSMOSIGGTGTLWPCRDRPSALHPTSVHARLPADIQWALGD-----LTAGNG 753
Db          546 APEGLVFAGLM--GLYDPPRDPVRA-----IRRUTTGGRVVMITGDSATAALSIGR 597
Oy          754 IGSKEPDDLRDVTTQYRGQLSYASGGDSLEN-----VTLPL---NIRPEPNRNTGY 801
Db          598 IGMPL--MFGTOSVTEGSKLMSDQALDECLOTAISIARTPSPEPKIVKGFORR--GD 653
Oy          802 AVG-TGD-ANDTNAFLQAVPGAKEDL--MSQVTLMOOKDORVNFRHDWKVITL 856
Db          654 VVAMTGDGVND----APALKLADIGTAMGCGTDAKEAAMILMTDDPATILSAI 705
Oy          857 IGCSALCDYCTDSNLNSLYAANFVD-HURNALDVLRHEPVRYLVNLDPFLNP--TIMQVFL 913
Db          706 EBGK3IFN-----NIRNFITFQOLSTSMAALSIIVAVATINGLENPLNPWQILWINILM 757
Oy          914 GNDDKCPCIQQASV 926
Db          758 DGP---PAQSLGV 767

RESULT 9
MNN4_YEAST      STANDARD;          PRT: 1178 AA.
ID  MNN4_YEAST
AC  P36044; P36043; P89095;
DT  01-JUN-1994 (Rel. 29, Created)
DT  15-DEC-1998 (Rel. 37, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DB  MNN4 protein.
GN  MNN4 OR YKL200C/YKL201C.
OS  Saccharomyces cerevisiae (Baker's yeast).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycetidae;
OC  Saccharomycetales; Saccharomycetaceae; Saccharomycetes;
OC  NCBI_TaxID=4932;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=S288C;
RX  MEDLINE=97115967; PubMed=9023541;
RA  Odani T., Shimma Y.-I., Tanaka A., Jigami Y.;
RT  "Cloning and analysis of the MNN4 gene required for phosphorylation
of N-linked oligosaccharides in Saccharomyces cerevisiae.";  

RT  Glycobiology 6:805-810(1996).
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Maia e Silva A., Bossier P., Viluela C., Fernandes L., Soares H.,
RA  Guerreiro P., Rodrigues-Pousada C.,  

RL  Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC  -I- FUNCTION: MAY FUNCTION AS A POSITIVE REGULATOR FOR
    MANNOSYLPHOSPHATE TRANSFERASE. IS REQUIRED TO MEDIATE
    MANNOSYLPHOSPHATE TRANSFER IN BOTH THE CORE AND OUTER CHAIN
    PORTIONS OF N-LINKED OLIGOSACCHARIDES.
CC  -I- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
CC  -I- SIMILARITY: TO YEAST YKL061W
CC  -I- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
    FRAMESHIFTS, ONE OF WHICH PRODUCES TWO SEPARATE ORFS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```


| RESULT 11 | | | | | | |
|-----------|---|-----------|------|----------|--|--|
| AC | GCN1_YEAST | STANDARD; | PRT; | 2672 AA. | | |
| AC | P3382; | | | | | |
| AC | P3382; | | | | | |
| DT | 01-FEB-1994 (Rel. 28, Created) | | | | | |
| DT | 01-FEB-1994 (Rel. 28, Last sequence update) | | | | | |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update) | | | | | |
| DE | Translational activator GCN1. | | | | | |
| GN | GCN1 OR YGL195W OR G1118. | | | | | |
| OS | Saccharomyces cerevisiae (Baker's yeast). | | | | | |
| OC | Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; | | | | | |
| OC | Saccharomycetales; Saccharomycetaceae; Saccharomyces. | | | | | |
| OX | NBGI_TaxID=4932; | | | | | |
| RN | [1] | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | |
| RX | MEDLINE=93260304; PubMed=8497269; | | | | | |
| RA | Marton M.J., Crouch D., Hinnebusch A.G.; | | | | | |
| RT | "GCN1, a translational activator of GCN4 in <i>Saccharomyces cerevisiae</i> , is required for phosphorylation of eukaryotic translation initiation factor 2 by protein kinase GCN2," | | | | | |
| RT | Mol. Cell. Biol. 13:3541-3556(1993). | | | | | |
| RN | [2] | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | |
| RC | STRAINS=S288C / FY1679; | | | | | |
| RX | MEDLINE=9119791; PubMed=9046087; | | | | | |
| RA | Cogliorni M., Klima R., Bertani I., Delneri D., Zaccaria P., Bruschi C.V.; | | | | | |
| RT | "Sequencing of a 40.5 kb fragment located on the left arm of chromosome VII from <i>Saccharomyces cerevisiae</i> ," | | | | | |
| RL | Yeast 13:55-64 (1997). | | | | | |
| CC | -1 FUNCTION: TRANSLATIONAL ACTIVATOR OF GCN4. MAY BE INVOLVED IN SENSING CHARGED TRNA AND STIMULATING THE KINASE ACTIVITY OF GCN2 IN AMINO ACID-STARVED CELLS. REQUIRED IN VIVO FOR THE PHOSPHORYLATION OF EIF-2-ALPHA ON SERINE-52 BY THE PROTEIN KINASE GCN2. | | | | | |
| CC | -- SUBUNIT: COMPONENT OF A HETEROGENIC COMPLEX THAT INCLUDES GCN1 AND GCN2. | | | | | |
| CC | -- SIMILARITY: STRONG, TO S.POMBE SPAC18G6_05C. | | | | | |
| CC | -- SIMILARITY: Contains 20 HEAT repeats. | | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | | | | | |
| CC | -- EMBL: L12467; AAA34635; 1; -; | | | | | |
| DR | EMBL: X91837; CA62949; 1; -; | | | | | |
| DR | EMBL: 272717; CAM96907; 1; -; | | | | | |
| DR | PIR: A49126; A4816. | | | | | |
| DR | SGD: SGD3153; GCN1. | | | | | |
| DR | GO; GO:0005330; C-cytosolic ribosome (sensu Eukarya); IDA. | | | | | |
| DR | GO; GO:006548; Pre-gylation of translational elongation; IMP. | | | | | |
| DR | InterPro: IPR001537; HEAT repeat. | | | | | |
| DR | PROSITE; PS00077; HEAT_REPEAT; 4. | | | | | |
| KW | Translation regulation; Activator; Repeat. | | | | | |
| FT | REPEAT 932 970 HEAT 1. | | | | | |
| FT | REPEAT 1030 1067 HEAT 2. | | | | | |
| FT | REPEAT 1099 1138 HEAT 3. | | | | | |
| FT | REPEAT 1243 1281 HEAT 4. | | | | | |
| FT | REPEAT 1284 1321 HEAT 5. | | | | | |
| FT | REPEAT 1405 1442 HEAT 6. | | | | | |
| FT | REPEAT 1523 1559 HEAT 7. | | | | | |
| FT | REPEAT 1561 1598 HEAT 8. | | | | | |
| FT | REPEAT 1603 1640 HEAT 9. | | | | | |
| FT | REPEAT 1641 1679 HEAT 10. | | | | | |
| FT | REPEAT 2290 2328 HEAT 11. | | | | | |
| FT | REPEAT 2347 2384 HEAT 20. | | | | | |
| FT | REPEAT 2672 AA; 296693 MW; 980FDD03753E9D1C CRC64; | | | | | |
| FT | SEQUENCE 213; Conservative 139; Mismatches 339; Indels 381; Gaps 53; Matches 213; Best local Similarity 19.9%; Score 123.5; DB 1; Length 2672; SQ | | | | | |
| QY | 165 KLUINVFFSN-----ASOCCYCPSEAOQNGLAGSYDELMLGVLDYLOQEVP 208 | | | | | |
| Db | 1447 KLIPIAVSNLDEIARWTRKRGSVQLIGNMAYLDP-OLSASLSTIPEIVGLVNDSHKEVR 1505 | | | | | |
| QY | 209 RAFAVLVDLSSEVARUSRQHIGTWLSPAPPCNGSEETTLAKVUMQWSQBAWNSLASS 268 | | | | | |
| QY | 1506 KAAB-B-SLKPGFGVIR-----NPEIQKLVPVILQ-----AIGDPT 1539 | | | | | |
| QY | 269 RYSQESFTVVFQDFP-FYENTPSLHSEDPRQDSTTLAWLWNRMMEPAGEKDPRLPSVKG 327 | | | | | |
| Db | 1540 KTYEALDLSLIQITQVHIDGPSL-----ALIIHHR-----GMHDSDANIKR 1583 | | | | | |
| QY | 328 GRPKMKPSQPSVPLFSYRSYNNYLRLQKQPKQKLEVER-----GAEI-RCP 371 | | | | | |
| Db | 1584 ---RACKITVGMAILVTDKDLIPYQQLIDEVELAMVDPVPVNTRATAALGAVALVERLG 1639 | | | | | |
| QY | 372 DKDPSDTVTISVHLKPADINVAGLGSQLTAGAGSPTPGNVLDVLTORGWSWVG 430 | | | | | |
| Db | 1640 EEQFDPLIP-----RLDUTLSDSESKSDRIGSA-----QALAEVIGL 1677 | | | | | |
| QY | 431 ---DENITGT-----VTTIANLILB-FNFSLKGFSVGTGKETPSNAFLN---QAVAGGRA 477 | | | | | |
| Db | 1736 DNDENIRDTALKAGKLUVKNYATKAVDVLILPELERGMEDNDRIRLSSVOLTGELLFQVT 1795 | | | | | |
| QY | 1678 GLTKLDEMELPTILAGVNFVRAKXREGFMFLLFPVCFQSOAP-YINQIQLFISGLA 1735 | | | | | |
| QY | 478 E-----DLPVQARL-----VPLM-----KND-----494 | | | | | |
| Db | 1796 GISSRNEFSE-----DGDRGEFSKLVDPVLQDRDRILLAFLVCRNDTSGIVRA 1847 | | | | | |
| QY | 539 ---DTIHAEV-----PRAFVNLUVTBLIUNRLRELYQEKKVYCPRMILRSCLCPVCKLKDNN- 591 | | | | | |
| Db | 1848 TTVDIWKALVNPTRAVKELPILPTGMIWTHLASSNV-LRNIAQTLGDLVRVGGNAL 1906 | | | | | |
| QY | 592 TELATLIEFKKPFQKTHOLIESQSPYDTPFVVWQPFENVUMPKT-----641 | | | | | |
| Db | 1907 SQLPSLE-----ESLIEITSNSDSDROQCVIALVLTIESASSTETTSQFOSTVNIIR 1957 | | | | | |
| QY | 642 GLPINSFFAPDCFHPSKSHRSRASLMNMFLPGQKTRHAFENKTCNQVOPF 700 | | | | | |
| Db | 1958 TALIDES-----ATVREAAALSFDV-----FQDVUGKTADEVLPY 1993 | | | | | |
| QY | 701 LRTYKNSMOHGHTWLPCRDRAPSALHPTSVHALRADIQVVAALGDSLTAGGNGSKPD 760 | | | | | |
| Db | 1994 L-----LHMLESSNDSPALIG-----LOEIMSKSD 2020 | | | | | |
| QY | 761 L-----PDTTQYRGLSYASAGGCSLEN-----TTLPNLREFRNNTGYAVGTFGD 807 | | | | | |
| Db | 2021 VIFPILPITLAPPID-----AFRASALGSAVAGLAVALKRSITINALVDAIIGTSE 2074 | | | | | |
| QY | 808 ANDTMFLNOAVPGA----KAEDMSQVTLKMKDH----RV-----NFRBDWKV 852 | | | | | |
| Db | 2075 DESTKGALBLALDRVFLSTNDDEGHPLIQQMILKSNEIKEIAVLERLPLPNFDDK-- 2131 | | | | | |
| QY | 853 TVLJGGSDLCYCDTSNLYSAAMFVDFHLRNALOVHREVRVPLVNLVFLNPITMRQVF 912 | | | | | |

| | | | | |
|--|---|---------------|--|--|
| Db | 2132 -TVL-----DFWVY-IPNFVSHAILSD---DEQDQVNGNENALS-TLKKVD | 2174 | CC | This SWISS-PROT entry is copyright. It is produced through a collaboration |
| Oy | 913 LGNDKC--PVQQRASVLC-----NCVHTL-----RENSOBLA | 942 | CC | between: the Swiss Institute of Bioinformatics and the EMBL outstation - |
| Db | 2175 KPTILEKLVYKPKAKOSLALTGQGDVAFKLPRGPNCVLPFHLGLMGNSNDREES-ALA | 2233 | CC | the European Bioinformatics Institute. There are no restrictions on its |
| Oy | 943 RLEAFSRAYRSURELUGSGRVTQEDPSWVQPFQNIQLPVLADGUDPDISFFA----- | 997 | CC | use by non-profit institutions as long as its content is in no way |
| Db | 2234 IADVSKTKPAALNPKPFS---VITGPLRVUGERFSSDIKATL-----FALNVLF | 2281 | CC | modified and this statement is not removed. Usage by and for commercial |
| Oy | 998 --PDCHPNQKHSQALARALTMLELGSKETLDRAEMPTCTQNEP 1046 | | CC | entities requires a license agreement (see http://www.isb-sib.ch/announce/ |
| Db | 2282 IKIPMFLRP--FIPQLORTFWVSLSD---ATNETIRLRAAKALGALIEHQ 2327 | | CC | or send an email to license@isb-sib.ch). |
| RESULT 12 | | | CC | EMBL; AC011620; AAF26134.1; - |
| ID PDP2_ARATH | STANDARD; | PRT; 1039 AA. | DR | InterPro; IPR001849; PH. |
| AC Q9M9WB; | | | DR | InterPro; IPR001683; PX. |
| DT 28-FEB-2003 (Rel. 41, Created) | | | DR | InterPro; IPR00169; PH; 1. |
| DT 28-FEB-2003 (Rel. 41, Last annotation update) | | | DR | InterPro; IPR00164; PLDC; 2. |
| DE Phospholipase D p2 (EC 3.1.4.41) (AtPdp2) (Phospholipase D2 PHOX and | | | DR | SMART; SM0023; PH; 1. |
| DE PK containing domain) (Phospholipase D zeta 2) (Pluzeta2). | | | DR | SMART; SM00155; PLDC; 2. |
| GN PLDP2 OR At3G05330 OR F18C10. | | | DR | PROSITE; PS50003; PH DOMAIN; 1. |
| OS Arabidopsis thaliana (Mouse-ear cress). | | | DR | PROSITE; PS50035; PLD; 2. |
| OC Buxariotae; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina; | | | DR | PROSITE; PS50195; PX; FALSE NEG. |
| OC Spermatophytina; Magnoliophytina; eudicots; Rosidae; | | | FT | Kinase; Hydrolase; Lipid degradation; Repeat; Multigene family. |
| OC euroids IT; Brassicales; Brassicaceae; Arabidopsis. | | | FT | DOMAIN 45 205 PH. |
| OX NCBI_TAXID:3702; | | | FT | DOMAIN 215 343 PH. |
| RN [1] | | | FT | DOMAIN 465 492 PLD PHOSPHODIESTERASE 1. |
| RP SEQUENCE FROM N.A. | | | FT | DOMAIN 840 867 PLD PHOSPHODIESTERASE 2. |
| RC STRAIN=CV. Columbia; | | | FT | ACT SITE 470 470 POTENTIAL. |
| RX MEDLINE=21016720; PubMed=11130713; | | | FT | ACT SITE 472 472 POTENTIAL. |
| RA Salanoubat M., Lemcke K., Rieger M., Ansorge W., Unseld M., Obermaier B., | | | FT | ACT SITE 477 477 POTENTIAL. |
| RA Partmann B., Vaille G., Bloecker H., Perez-Alonso M., Obermaier B., | | | FT | ACT SITE 845 845 POTENTIAL. |
| RA Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P., | | | FT | ACT SITE 847 847 POTENTIAL. |
| RA De Simone V., Choisne N., Artiguenave F., Robert C., Brötzner P., | | | FT | ACT SITE 852 852 POTENTIAL. |
| RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Quétier F., | | | SQ | SEQUENCE 1039 AA; 117902 MW; AB433C23T3B77F3 CRC64; |
| RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Bens V., | | | Query Match 1.6%; Score 123; DB 1; Length 1039; | |
| RA Wurmback E., Drzencik H., Erflie H., Jordan N., Bangert S., | | | Best Local Similarity 18.7%; Pred. No. 5.4; Gap 37; | |
| RA Wiedemann R., Kranz H., Voss H., Holland R., Brandst P., Nyakatura G., | | | Matches 157; Conservative 123; Mismatches 313; Indels 246; Gap 37; | |
| RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B., | | | Oy 412 GVNLDVLT-QVRGLSWSVGGBDENIGVTTLANILREFNPLKGFSVGTGKET----- 462 | |
| RA Conradi A., Hornischer K., Kauer G., Loehnert T.-H., Nordsieck G., | | | Db 273 GKUDINMVFDLGLQ---GKXESSRPRPLAQVKHNPRLRGFKTSGDRTRVRRTSS 328 | |
| RA Reichelt J., Scharfe M., Schoen O., Bargnes M., Terol J., Clement J., | | | Oy 463 -----SPNAPLANOQAVAGRAEADLPYQARRND-----LMKNRDT 496 | |
| RA Navarro P., Collado C., Perez-Perez A., Otterwaelder B., Duchemin D., | | | Db 329 RKVKEWVKAVIDAGCVSPHRGSFAPRGLTS-DSQAOQFWGHTAFFATAIQNATS 387 | |
| RA Cooke R., Laude M., Berger-Lacroix C., Purnelle B., Masly D., | | | Oy 497 IHRQEDKMKITLFIGNDLCDPCNDLVHISQNQFTNIGKALDILHABVPRAFVNJVTL 556 | |
| RA de Haan M., Maatz A.C., Alcaraz J.-P., Cottet A., Casacuberta E., | | | Db 388 ELPMTGWL-----CPEL-YLRKPPEHPSLPSLALIETAKOGKVQOI- 431 | |
| RA Monfort A., Argirou A., Flores M., Ligouri R., Viala D., | | | Oy 557 ETVNRLRELYQEKKVYCPRMILSRLCPVLFDDNSTBLATLJEFNKKPQEKTQH----- 611 | |
| RA Mannhaupt G., Haese D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W., | | | Db 432 -ALKINSLYSKGRQLQINHKAVK---VLRVPDHLSGGIYLMWHHEKIVTDYQCFIGG 485 | |
| RA Mayer K.F.X., Kaul S., Town S., Koo H.-L., Tallon L.J., Jenkins J., | | | Oy 612 -IESGRYDTEBEDFTWVQPF--ENVDMKTSBGLDPSRFAPCFCHESSKSHR--- 663 | |
| RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P., | | | Db 486 IDLCFGRYDTAIIHGKIDGCCPPYVTPGKQYVNPRESB---PNMSBETMKDLDKRYPRW 542 | |
| RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S., | | | Oy 664 -AASAIWNMLPEVGOKTRHKPFENKINTCPNQVQPFURTYKNSMQHGKHTWLPCDRA 721 | |
| RA Pai G., Mittscher J., Sellers P., Gill J.E., Feldlum T.V., | | | Db 543 HDVCAWGPCCRDVARYRHFVORWNHKSRRNKAPNEQTIPPLMPHHMLPH--YLGRF-- 598 | |
| RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C., | | | Oy 722 PSALHPTSVHALRP--ADIQVAAALGDSLTAQNGIGSKPDOLPDVTTQYRGLSYAGGD 778 | |
| RA Fraser C.M., Kaneko Y., Nakamura Y., Sato S., Kato T., Asamizu E., | | | Db 599 ----IDIAAAKPEPDPKVFLRHDSSS---ASPPQEIPLPQETDADFGRGD 649 | |
| RA Sasamoto S., Kimura T., Idesawa K., Kawashita Y., Kishida Y., | | | Oy 779 -----GSLENVTNTLRLR-----NRNLIGYAVGTGDANDT-- 811 | |
| RA Kyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A., | | | Db 650 LKLDGARRQDGGETSEBDLDAVNWWQIGKQSDCCRQCTRSVSOWAGTSQPDSIH 709 | |
| RA Nakayama S., Nakazaki M., Shimo S., Takeuchi C., Wada T., | | | Oy 812 -----NAFLNQAVGKAED--LMSQVQTLQMRKDDHVNRFHDWKV 852 | |
| RA Watanabe A., Yamada M., Yasuda M., Tabata S.; | | | Db 710 RAYCSLQIQAHRPIYIENOFF---ISGLEKDTIILVRBALYRI----LKANEENKC 761 | |
| RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis thaliana." | | | | |
| RL Nature 408:820-822 (2000). | | | | |
| CC -- FUNCTION: Hydrolyzes glycerol-phospholipids at the terminal phosphodiesteric bond. Phosphatidylcholine + H(2)O = choline + a | | | | |
| CC -- CATALYTIC ACTIVITY: A phosphatidate phosphotidate. | | | | |
| CC -- ENZYME REGULATION: Calcium-independent and PIP2-dependent. | | | | |
| CC -- SIMILARITY: Belongs to the phospholipase D family. PXPH-P LD subfamily. | | | | |
| CC -- SIMILARITY: Contains 1 phox homology (PX) domain. | | | | |
| CC -- SIMILARITY: Contains 1 PH domain. | | | | |
| CC -- SIMILARITY: Contains 2 PLD phosphodiesterase domains. | | | | |

QY 853 ITVLI-----GGSDLCDYCTDSNLYSAMFVDHJRNALDVHLREVPRVLVNLDFL 903
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 762 FTVWIVPLPGFGGGIN-----DFGA-----VTALEMWQYRTSRSGTSILDNL 808
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 QY 904 N-----PTIMROVFLGN-----FDKCPVQQSVLCNCVL 932
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 809 NALIGPKTQDYISFYGLASYGRIFEDGPIATSQIYVHSKLMIVDDRAVIGSSNINDSL 868
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 QY 933 TLRBNSOBLARLE--AFSRAYRSRMRELVGSGRYDQ----EDFSVWVQPFQFNIOLPV- 985
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 869 LGSRDSEIGVWIEDKEFVESSNGMKWAGKFSYLSRCSLWISHLGHAGEICKIEDPIK 928
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 QY 986 -----LADGUDTSPAP--DCIHONQKFISQLAGALWML---EPISSKTETDIL 1032
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 929 DATYKDLWNMATAKNTDYNQVFSCI-EINE-HIRSAIRHMALCKDKLQHHTIDGI 985
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 QY 1033 RAEMPITCPTQNPBPFLTRPRNSNTYPRPALENWGSDFLCTEMKASNSVPPSVHQRP 1091
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 986 APERLESCSDSWELKETRGNLVCFPLQ-----FMCDQ-----EDURP 1024
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 RESULT 13
 GALX_SCBPO STANDARD; PRY; 713 AA.
 ID GALX_SCBPO
 AC QPHD03;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE GAL10 bifunctional protein [Includes: UDP-glucose 4-epimerase
 DE (Mutarotase)].
 DE (Galactowaldenase); Aldose 1-epimerase (EC 5.1.3.3)
 DE (Mutarotase).
 GS GAL10 OR SPBpB2B2.12C.
 OS Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TAXID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21868401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle B.J., Hunt S., Jagels K.,
 RA Jones K., Jones L., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitzch E.,
 RA Rutherford S., Saunders D., Seeger K., Sharp S.,
 RA Shelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vansreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabell C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Rohl T.M.,
 RA Eber P., Zimmermann W., Wedler H., Wambutt R., Pur nelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lejaire V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Dabag R.R., Cruzado L., Jimenez J., Sanchez J., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Folsburg S.L.,
 RA Carrutti L., Lowe T., McCombi W.R., Paulsen I., Potashkin J.,
 RA Shapovskii G.V., Usbery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -I- FUNCTION: Mutarotase converts alpha-aldoose to the beta-anomer. It
 is active on D-glucose, L-arabinose, D-xylene, D-galactose,
 malto- and lactose (By similarity).
 CC -I- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -I- CATALYTIC ACTIVITY: Alpha-D-glucose = beta-D-glucose.
 CC -I- COFACTOR: NAD.

CC -I- PATHWAY: Galactose metabolism; third step.
 CC -I- PATHWAY: Hexose metabolism.
 CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE SUGAR
 EPIMERASE FAMILY.
 CC -I- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ALDOSE
 EPIMERASE FAMILY.
 CC ---
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC ---
 DR EMBL; AL12522; CAC21414.1; -.
 DR HSSP; D09147; IAKL.
 DR InterPro; IPR01823; Ald1_Epimerase.
 DR InterPro; IPR001509; Epimerase_Dh.
 DR InterPro; IPR00586; Gale.
 DR InterPro; IPR00205; NAD binding.
 DR Pfam; PF01263; Aldose_epim_1.
 DR Pfam; PF01370; Epimerase_1.
 DR TIGR01179; Gale; 1.
 DR PROSITE; PS00545; ALDOSE_1_EPIMERASE_1.
 KW Multifunctional enzyme; Isomerase; NAD; Galactowaldenase.
 FT DOMAIN 1 350 GALACTOWALDENASE.
 FT DOMAIN 351 713 MUTAROTASE.
 FT NP BIND 37 38 NAD (POTENTIAL).
 FT ACT SITE 532 532 MUTAROTASE (POTENTIAL).
 SQ SEQUENCE 713 AA; 80666 MW; 51C89DA0843A8556 CRC64;
 Query Match 1.6%; Score 12.5%; DB 1; Length 713;
 Best Local Similarity 19.8%; Pred. No. 3.2; Mismatches 235; Indels 231; Gaps 38;
 Matches 138; Conservative 92; Mismatches 235; Indels 231; Gaps 38;
 QY 370 CPDKDPSD-----TVFTSVHLKPD-----INVIGALGDSTAGNGAGSTPG 412
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 149 CPN-DPTNPYKGKVKYAEENIKDQLTSNDNTMRGAILYFNPIGAHISGLIGEVDPL-I-PN 206
 QY 413 NVLDVLTQY--RGLLSWVGQDDE--NIGT-----VTTLA-----NLTREFNS-- 450
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 207 NLNPFLQAGVAIGREKLUVFGDHDGTPYRDYHVVDLAKGHIAALNTINKINSEG 266
 QY 451 -LKGSFSTGTGKESP---NATLNQAVAGGAQRAEDLP--VQARR--LVDLMNDTRHFO 500
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 267 MYREWNLGTGKESVSPDVIYHAKCEV---GKDLPVEVERGERTGDTNLNTASPRNASE 321
 QY 501 EDWKIIITLFIGGNDLDCFENDLVHYSQONFTDNIKGALDILHAEVPRAFVNLTVELVN 560
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 322 LKKK-----AELSIITPCARDLKWTEEN-----BPGF-----QIDN 352
 QY 561 LR-ELYDKKVYCPRMILRLSCPCVLUKDNNSTELATIENKKFQKTHQL---ESG 615
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 353 YKWLKFLNTLGIMDKYKNLHTIC-----FQDLEVSIANYGALVQAVRYKGRNLVNGFDS 407
 QY 616 RYDTREPTVVQOPPFENDMKTSEGIPDNSSFFAPCFHSSKSRRASALWNMLEP 675
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 408 RYKLKEN-----PFP-----GATGCRFAIRIANGOFEV 435
 QY 676 VGGTKTRKPE-NKINITCPNO-----VQPLRKY--NSMQ-----GHGTWLPCL 718
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 436 DGHLYTICKNEENKTTGCGNGFDKOFFLGLARQDYNLETFILVKOGNNGF--- 491
 QY 719 DRAPSALIPTSHALRDPDIOVVAALGDSLITAGNGLGSKPPDLPVTOYQGLSYSAGGD 778
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 492 ---PSDIEFLVVKYTKNSLEI-----BYKSVIPEYSKL----- 522
 QY 779 GSLENVTTLPNLREFRNRLGAV-----GTDGANDTNAFLN-QAVPGAKED 826
 :|: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |: |:
 DB 523 ---NVTVV-----NLTHSYNMLASPNTKPTGTTKSTTNVLYKVNSETSPLTG 569

RN 1063 SWKDVLLCISQSLRQLVLSAGVDINSPLDVSTTRPLRKSLDKNTQRSGSISLKHSKP 1122
 RP STRAIN=Berkley;
 RC MEDLINE=2019606; PubMed=10731132;
 RA Adams M.D., Celinkin S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scheerer S.E., Li P.W., Gallo R.F.,
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champé M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter B.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Balow R.M., Babu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botcham M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burton K.C., Busam D.A., Butler H., Cadie E., Center A., Chandra I.,
 RA Cherry J.M., Cauley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mayas A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doupe R., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin R., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann R.,
 RA Fosler C., Gabrilian A.E., Gang N.S., Gelbart W.M., Glasser R.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levittsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.N., Nelson D.L.,
 RA Nelson D.R., Nixon K., Nuskbner D.R., Pacieb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Strickss R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.-A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*";
 RT Science 287:2185-2195(2000).
 RN [3].
 RP FUNCTION.
 RX MEDLINE=91092252; PubMed=2124970;
 RA Keith F.J., Gay N.J.;
 RT "The *Drosophila* membrane receptor Toll can function to promote
 RT cellular adhesion";
 RL EMBO J 9:4299-4306(1990).
 CC -I- FUNCTION: REQUIRED FOR DORSAL-VENTRAL EMBRYONIC POLARITY. MAY
 CC FUNCTION AS A MEMBRANE RECEPTOR. PROMOTES HETEROPHILIC CELLULAR
 CC ADhesion.
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -I- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
 CC -I- SIMILARITY: Contains 21 leucine-rich (LRR) repeats.
 CC -I- SIMILARITY: Contains 1 TIR domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; ABB01758; AAF56624.1; -.
 DR EMBL; ABB01758; AAF56624.1; -.
 RP SEQUENCE FROM N.A.
 RP MEDLINE=88135760; PubMed=2449285;
 RA Hashimoto C., Hudson K.L., Anderson K.V.;
 RT "The Toll gene of *Drosophila*, required for dorsal-ventral embryonic
 OC polarity, appears to encode a transmembrane protein";
 OC Cell 52:269-279(1988).
 RL

DR GO; GO_000967; Panifungal polypeptide induction; IMP.

DR GO; GO_000950; Podosal/ventral axis specification; NAS.

DR InterPro; IPR004075; ILL_receptorl.

DR InterPro; IPR00483; LRR_Cterm.

DR InterPro; IPR00372; LRR_Nterm.

DR InterPro; IPR003591; LRR_TYP.

DR InterPro; IPR000157; TIR_domain.

DR Pfam; PF00460; LRR; 11.

DR Pfam; PF01463; LRRCT; 2.

DR Pfam; PF01462; LRRNT; 1.

DR Pfam; PF01582; TIR; 1.

DR PRINTS; PRO0019; LEUWICHRPT.

DR SMART; SMD0369; LRR_TYP; 3.

DR SMART; SMD0082; LRRCT; 2.

DR SMART; SMD0013; LRRNT; 1.

DR SMART; SMD0255; TIR; 1.

DR PROSITE; PS50104; TIR; 1.

DR PRINTS; PRO0019; LEUWICHRPT.

KW Developmental process.

KW Cell adhesion; Leucine-rich repeat; Signal.

FT SIGNAL 1 27 POTENTIAL.

FT CHAIN 28 107 TOLL PROTEIN.

FT DOMAIN 28 807 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 808 828 POTENTIAL.

FT DOMAIN 829 1097 CYTOPLASMIC (POTENTIAL).

FT REPEAT 149 172 LRR 1.

FT REPEAT 174 195 LRR 2.

FT REPEAT 196 213 LRR 3.

FT REPEAT 221 243 LRR 4.

FT REPEAT 245 267 LRR 5.

FT REPEAT 268 291 LRR 6.

FT REPEAT 293 317 LRR 7.

FT REPEAT 319 340 LRR 8.

FT REPEAT 341 364 LRR 9.

FT REPEAT 365 388 LRR 10.

FT REPEAT 390 412 LRR 11.

FT REPEAT 413 436 LRR 12.

FT REPEAT 438 459 LRR 13.

FT REPEAT 472 495 LRR 14.

FT REPEAT 496 521 LRR 15.

FT REPEAT 523 546 LRR 16.

FT REPEAT 550 574 LRR 17.

FT REPEAT 667 690 LRR 18.

FT REPEAT 692 713 LRR 19.

FT REPEAT 714 736 LRR 20.

FT REPEAT 785 807 LRR 21.

FT DOMAIN 857 996 TIR.

FT CARBOHYD 80 80 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 140 140 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 175 175 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 235 235 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 270 270 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 275 275 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 346 346 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 391 391 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 482 482 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 508 508 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 528 528 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 654 654 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 677 677 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 702 702 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 715 715 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 730 730 N-LINKED (GCNCAC. . .) (POTENTIAL).

FT CARBOHYD 738 738 N-LINKED (GCNCAC. . .) (POTENTIAL).

SQ SEQUENCE 1097 AA; 124655 MW; DIBFC42245E3BABE CRC64;

Db 228 LWSNQL---KHLTKHDFEGATSVLGIDDHNGIEOLPHDVFAHLTNVTIDINSANLFRS 283

Qy 704 YKNSMOHGHTWLP---CRDRAPSALHPTSHALRADIQVAALGDSLTAGNGIGSKPD 759

Db 284 LPQGLFPHNKHANEVLMNNVPLATLPSRAFLPANO---EQTLRLRB-----QSPLPG 335

Qy 760 DLPDVTTQYRGISYSAGGDSL-----NVTTLPNTLREFNRNL 798

Db 336 DPEHNSIQITWISL---GDNLIKTUPTTLEHQVNVLISLDLSNNRUTHLPSLPAFTNL 392

Qy 799 TCYAVG---TCGANDTNAFLNQAVIGAKA----- 825

Db 393 TDRLLEDNLNLTGISGDTSNGLGNLVLTIVMSRNRLRTIDSRAFVSTNGLRHLHLDNDL 452

Qy 826 ---DLMQVO-----TLMOKKODHRVNTHEDWKVUTVILIGGSPLCDYCTDSN 870

Db 453 QDPLMLQTOINSFGGMHGLTILNLNNISLYFDWKNMLORELOI-SYNNIS 511

Qy 871 1-YSAANFVDHURNALDVHLRBEVPRVIL---VNLFDPFLPTIMQVFLGNGPDKCPVQOAS 925

Db 563 IVCDCDTIWFIQULVRGVHKPQVSQRFKRTDLVCS-----QPN 601

Qy 986 LADGLPDTSSFFAPDCIHPNQKFSOLARALWTNMLBPLGSKTEFLDRAEMPTC -PTQ 1043

Db 643 DKALVINCHSGNLTH -VPRFLNKKOMQMLMELHENLTLLRPSANTPGESVSLHA 700

Qy 602 VIEGTPQR-----IEPP-----TLCPLDESDPRERKCPGNCNCFVTRY 642

Db 926 VLGNCVLTRENSQELARLEAFASRAVRSMSBLVGGRYDQEDESVLQPFONIOLPV 985

Db 701 GHNLTSDIVDQLPNTLHLDISW-----NHQLMNATVGLPFLT 740

Qy 1101 GDSLTTA-VGARPNNSSDLPSPWWSGCGDGNLTHTCLPNTIKKENYLLGF--ST 1157

Db 1158 SWEGTAGLNVAEG-----ARADMPAQADLVERMMKNSPD--INLEKDWKLVTLPFI 1208

Db 741 MKOR-----SVKLGPMCDTAKLFLQDNEFRIGENEMMCVNAEMPTMVELL-- 793

Qy 1209 GYNDLCHYCENPENAEAHATEYYOHIOQALDISELSELPRAVNVVEMBLASLYOGOGCKA 1268

Db 794 STNDIC-----PAEKGVFIALAVIALTGLAGFT-----AALYKFOTEK 835

Qy 1269 MIAAQNCTCLRHQSOSLEKBLKKUNMQHGISSFSYMQYQTEDPAVWVQPFFONT 1328

Db 836 IWLAYANLILWVFTEBDLKD-----KFDARFISYSHKD----QSPIEDY 876

Qy 1329 LTPLNERGDTLTFSEDFCFHFSD---RGHAEAMALWNNNLPV--GRKT--TSNNFT 1380

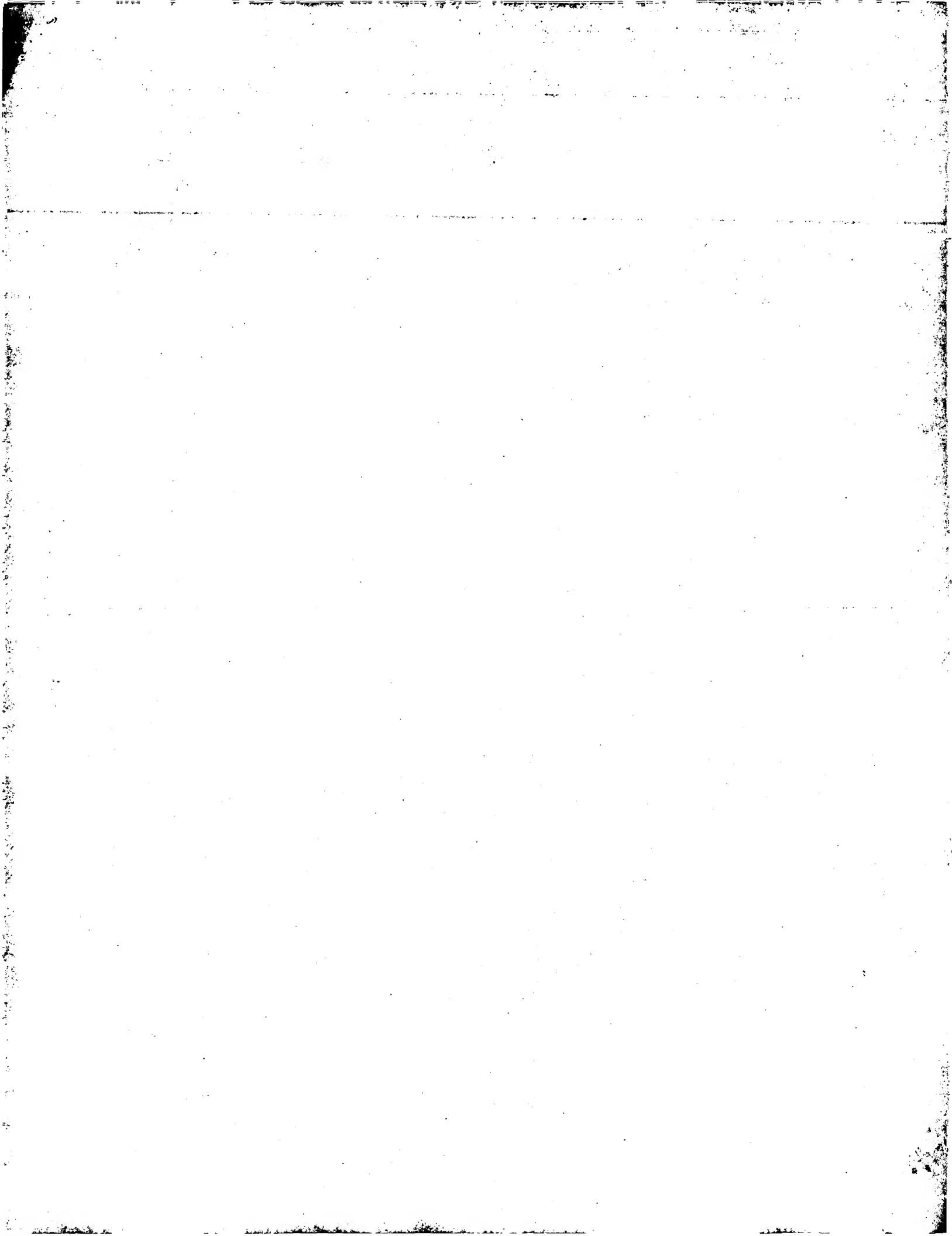
Db 877 LYQLEIGQPKFQL---CVHERDWLNGH-----IPENITMSVADSRRTIVSQFI 926

Qy 1381 HSR-AKUK 1387

Db 927 KSEWARUE 934

Search completed: January 6, 2004, 19:07:24
Job time : 44 secs

Query Match 1.6%; Score 121.5; DB 1; Length 1097;
Best Local Similarity 18.3%; Pred. No. 74; Mismatches 315; Indels 269; Gaps 37;
Matches 155; Conservative 109; Mismatches 315; Indels 269; Gaps 37;
Qy 668 LWNMMLSPVGQKTRHKEF---NKINITCPQVQPLRT 703



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

January 6, 2004, 14:46:04 ; Search time 44 Seconds

Title: Perfect score: .7766
Sequence: 1 MGIRPGIFFLBLLLLQGQT..... RCRRGGRRDPPMSLRTVAL 1458

Scoring table: BLOSUM62 , Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Match Length DB ID

Description

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|--------|--|
| 1 | 5987.5 | 77.1 | 1458 | 2 | A45665 | adult-specific brush border esterase/phospholipase (EC 3. - . -) precursor - rabbit |
| 2 | 613.5 | 7.9 | 382 | 2 | T20655 | hypothetical prote |
| 3 | 605.5 | 7.8 | 414 | 2 | T20833 | hypothetical prote |
| 4 | 463.0 | 6.0 | 981 | 2 | T1060 | hypothetical prote |
| 5 | 406.5 | 5.2 | 348 | 2 | T2016 | hypothetical prote |
| 6 | 347.5 | 4.5 | 349 | 2 | T2015 | hypothetical prote |
| 7 | 338.0 | 4.4 | 425 | 2 | T20835 | hypothetical prote |
| 8 | 159.5 | 2.1 | 941 | 2 | F97353 | uncharacterized co |
| 9 | 149.0 | 1.9 | 3655 | 2 | T30804 | TRAP-like protein |
| 10 | 146.5 | 1.9 | 1755 | 2 | F82618 | chemoraxis-related |
| 11 | 141.5 | 1.8 | 913 | 2 | T15278 | hypothetical prote |
| 12 | 133.0 | 1.7 | 1708 | 2 | F6189 | protoporphyrin IX |
| 13 | 131.5 | 1.7 | 1361 | 2 | T30884 | neural specific DN |
| 14 | 130.5 | 1.7 | 891 | 2 | E95590 | hypothetical prote |
| 15 | 130.5 | 1.7 | 1165 | 2 | T35279 | hypothetical prote |
| 16 | 130.5 | 1.7 | 2471 | 2 | T24589 | hypothetical prote |
| 17 | 129.5 | 1.7 | 1245 | 2 | T42920 | hypothetical prote |
| 18 | 128.0 | 1.6 | 589 | 2 | G8650 | polysaccharide bio |
| 19 | 127.5 | 1.6 | 1016 | 2 | T30553 | disease resistance |
| 20 | 127.5 | 1.6 | 1280 | 2 | T33357 | hypothetical prote |
| 21 | 126.5 | 1.6 | 1498 | 2 | E84302 | hypothetical prote |
| 22 | 126.5 | 1.6 | 2471 | 2 | T42977 | large tegument pro |
| 23 | 126.0 | 1.6 | 1028 | 2 | I5173 | myosin I beta - bu |
| 24 | 125.5 | 1.6 | 1178 | 2 | S7475 | mannosylphosphoryl |
| 25 | 125.0 | 1.6 | 1650 | 2 | S26721 | hypothetical prote |
| 26 | 124.5 | 1.6 | 1946 | 2 | AC2141 | serine/threonine k |
| 27 | 123.5 | 1.6 | 1230 | 2 | E64664 | outer membrane pro |
| 28 | 123.5 | 1.6 | 1534 | 2 | G91017 | probable membrane |
| 29 | 123.5 | 1.6 | 1534 | 2 | A85862 | hypothetical prote |

ALIGNMENTS

RESULT 1
A45665

C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 08-Oct-1999

C;Accession: A45665
R;Boll, W.; Schmid-Chanda, T.; Semenza, G.; Mantei, N.

J; Biol. Chem. 268, 12901-12911, 1993

A;Title: Messenger RNAs expressed in intestine of adult but not baby rabbits. Isolation & characterization of cDNA clones

A;Reference number: A45665; MUID:93286138; PMID:8509424

A;Status: Preliminary

A;Molecule type: mRNA

A;Residues: 1-458 <BCL>

A;Cross-references: GB:Z12841; NID:91689; PID:CAA78303.1; PID:91690

C;Keywords: hydrolase

Query Match Best Local Similarity 77.1%; Score 5987.5; DB 2; Length 1458;
Matches 1130; Conservative 127; Mismatches 197; Index 3; Gaps 3;

Ory 1 MGIRPGIFFLBLLLLQGQTPOINTSPRKSTLEPLQWPTLKNSPPCPNPNLKGUNPKSK 60

Db 1 MAWPSPVFLGLPLGLPGRGADQIQTSQSGKNTLGQWLWPESLKTFFPPCDPKTLAESVE 60

Ory 61 SPSHLSPDIKPIKVAAGNLEPPDPETGDLKDQDWIE-RPQVCMGUMTUDIYFSP 119

Db 61 SVHSLRSPDIKPIKVAAGNVEPATPDGADLDRBDQTEKRPQAOCMGVNTVLSIDIGRFP 120

Ory 120 SVPMPVCHTGKEVIPHDGAELWIQAQELVRANKENKLQDFDWKLINIFNSNASOCYL 179

Db 121 SALMPIC-PETRLVPIGGADLWMQATEVSMRNLQDEPHDWKLINIFNSNTSQC 179

Ory 180 CPSAQOQNGLAAGGVDELMGVLDQEVPAVFLVYDSEVAEVSROQYHGTWLSAPEPC 239

Db 180 CPSAQOKGLVLGGMQDKLTTRDLYQEVPKAFVNLDSELAAFSRWQGAOLSPAAEPC 239

Ory 240 NOSEETTRALKUMQSYQEWNLSLASRSVSEQSEFTWQTPFETTSLHSEPRQLQ 299

Db 240 RCLRETSQTLKVTQTYLAWDSLASKNTQSFAVVQPFYESSLSALLAEPLQ 299

Ory 300 DSTLAWHLKMRMMPAGEKOBPLSYKGRPMKCCSQBSPPLFSYKRNQNYJTRQPKDK 359

Db 300 DPTTLALSLWTRMMEDIGKRPESKERKLRCPTQDSEPLFTVNSGQLTROSPQK 359

Ory 360 LEVREGAETRCPDQSDTVTSVRLKPADINVAGLGSQSLTAGNGAGSPGNVLDLT 419

Db 360 LEVREGTEIRCPDKDPSDVSPTSVHLKPADIKVIGAMGPDSLTAQNGAGSOPGNILDT 419

Ory 420 QYRGLEWSVGQDENIGTVTIANIREBNFNSLKGIGSVGCKETSNAFLQAVAGRAE 479

translation activa
probable protein t
Toll protein precu
dynein heavy chain
extracellular neut
probable protein t
adenomatous Polypo
probable outer mem
serine proteinase
cyclosporin Synthe
probable swf/snf h
botulinum neurotox
hypothetical prote
hypothetical prote
hypothetical prote
ATM-like protein [
hypothetical prote

| | | | | | | |
|----|-------|-----|-------|---|--------|--------------------|
| 30 | 123.5 | 1.6 | 2672 | 2 | A8126 | translation activa |
| 31 | 122 | 1.6 | 1811 | 2 | T39252 | probable protein t |
| 32 | 121.5 | 1.6 | 1097 | 2 | A29943 | Toll protein precu |
| 33 | 121 | 1.6 | 4725 | 1 | A44357 | dynein heavy chain |
| 34 | 120.5 | 1.6 | 836 | 2 | D97182 | extracellular neut |
| 35 | 120.5 | 1.6 | 1570 | 2 | T38792 | probable protein t |
| 36 | 120.5 | 1.6 | 2843 | 1 | RIBHUP | adenomatous Polypo |
| 37 | 120 | 1.5 | 1237 | 2 | D71850 | probable outer mem |
| 38 | 120 | 1.5 | 2140 | 2 | F95074 | serine proteinase |
| 39 | 120 | 1.5 | 15281 | 2 | S41309 | cyclosporin Synthe |
| 40 | 119.5 | 1.5 | 1163 | 2 | E71481 | probable swf/snf h |
| 41 | 119 | 1.5 | 1276 | 2 | S11455 | botulinum neurotox |
| 42 | 119 | 1.5 | 1107 | 2 | T25450 | hypothetical prote |
| 43 | 119 | 1.5 | 1116 | 2 | T6112 | hypothetical prote |
| 44 | 119 | 1.5 | 2089 | 2 | C5426 | ATM-like protein [|
| 45 | 119 | 1.5 | 3738 | 2 | T05501 | hypothetical prote |

Db 420 QYRQLSWVGQDQNIStVTLANILREFNPSLOGSVGTGRETTSQAFNQAVAGARADG 479
 Qy 480 LPPVQARRLYDMDKNDTRHFOEDKMITLFIGNDLCDPNCNDLVHYSQNFTDNIGALD 539
 Db 480 LIPQAQRQLVALMNDTRINFOEDKMITLFIGNDLCDPNCNDPVRSQNFTDNIGALD 539
 Qy 540 ILLHEVPAFVNLTVEIUVNLRYQEKKVYCPRMILSUCPCVLUKDDNNSTELATLIE 599
 Db 540 ILLAEIPPAFVNLTVEIUVNLRYQEKKVYCPRMILSUCPCVLUKDDNNSTELATLIE 599
 Qy 600 FNKHQEQKTHQLLIESGRVDTREDTUVVQPFENNDMPKTSSEGPDNSFFAPDCFHSSK 659
 Db 600 TIKYQERTQQLIDSGRDUTRDPTVVLQPFERKUMRKTQDSIPDNSSFFAPDCFHSSK 659
 Qy 650 SHRAASALWNMNLLEPVGQKTHINDFEGLAVNTCPNOVWPFLSYKNSVQGEGFLWLPKD 719
 Db 650 AHAHAASALWNMNLLEPVGQKTHINDFEGLAVNTCPNOVWPFLSYKNSVQGEGFLWLPKD 719
 Qy 720 RASALHPTSVHAIRPADIQVVAALGDSITAGNIGSKSPDPLADVTQYRGISYASGD 779
 Db 720 RSASAPPTSVHAIRPADIQVVAALGDSITAGGIGSKENDSDGTQYRGISYSSGD 779
 Qy 780 SLENVTTLPNLAREFNRLTGAYVGTGDAANDTNAFLNQAVPGAKAEDIMSQVOTLWQMK 839
 Db 780 SLDNVTTLNLIROFNSNLMGFAVGTDGASGTFNAFFNOAVPGAKARDMSQVOTLWQMK 839
 Qy 840 DHARVNFBEDHWKVITVLGGSDACDYCDTSNLASAANFDHLANALDVHREVPVNL 899
 Db 840 DHDARVNFOEDHWKVITVLGGSDACDYCDTSNLASAANFDHLANALDVHREVPVNL 899
 Qy 900 VDELNPTIMQVLFGNPKCPOQAVSILCNCVULRENSQELARLEFSRAXRSMRBY 959
 Db 900 VDEKNPSVTRQVLFGNPKCPOQAVSILCNCVULRENSQELARLEFSRAXRSMRBY 959
 Qy 960 GSGRYDTBDFSVLQLQPFRONIOPVLAQDGLPDTFFAPDCIAPNQKHSQJARALWNM 1019
 Db 960 ESGYDVTRADFSVLTQPFPHSIQLPVLDGRUITSFAPDCVPHNPKHSQJARALWNM 1019
 Qy 1020 LEPGSKYETLDIRAEMPTCTONEPLRTPRNNSYTPKIAIENGSDFCTEWKA 1079
 Db 1020 LEPPLGKTDALDIDTAITLCTCPTONEPLRTPRNNSYTPSRAVENNGSDFCTAMMAS 1079
 Qy 1080 NSVPTSHOLRADIKVVALGSITAVGARNPSOLPSWSRGLSNSIGDIONEHT 1139
 Db 1080 RGVEVNSVHLOPEDIKVVALGSITLAMGARNPSNSDPPMFRGLSNSIGDGALEHT 1139
 Qy 1140 TLPMLKGENPKYLGFSTSTWEGTAGLNVAEGARADMPAQADLVERMKNSPDINLK 1199
 Db 1140 TLPMLKGENPKYLGFSTSTWEGTAGLNVAEGARADMPAQADLVERMKNSPDINLK 1199
 Qy 1200 DWKLVLTFIGNDLCHYCNPPEAHATEYVQHQAQDPLISESLPRAFYVNUVEMELA 1259
 Db 1200 DWKLVLTFIGNDLCHYCNPPEAHATEYVQHQAQDPLISESLPRAFYVNUVEMELA 1259
 Qy 1260 YQOGGKCA-MLAQNNCTCLRISSQSLKEQLEKVVNMLQHGCISSSYWMQTOREDFA 1318
 Db 1260 HODOGRCATLAAQSHCTCPYSSVEMQELKVVNMLQGSLRSYSHOTOVEDFA 1319
 Qy 1319 VVTOQPFQONTLTPNERGDTLTFSEDCFHFSDRGHAEMLAALWNMNLLEPVWKTNSN 1378
 Db 1320 VVQPFQONTLTPNERGDTLTFSEDCFHFSDRGHAEMLAALWNMNLLEPVWKTNSN 1379
 Qy 1379 FTHSRAKUKCSPSPSPYLYTNRSLRPLPOAERAPEVLYWAQVPAAGYLVGUGIYTWW 1438
 Db 1380 PTYSRTKLUCKSPSPSPYLYTNRSLRPLQEAADPTVLYWAQVPAAGAGLILAMVAG 1439
 Qy 1439 RCRRGGRRDPPMSLT 1455
 Db 1440 RGMCRCPREDPPLSLST 1456

hypothetical protein F09C8.1 - *Caenorhabditis elegans*
 C-Species: *Caenorhabditis elegans*
 C-Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
 C-Accesion: T20655
 R. McMurray, A.
 Submitted to the EMBL Data Library, November 1995
 A-Reference number: Z19306
 A-Accesion: T20655
 A-Status: preliminary; translated from GB/EMBL/DDJB
 A-Molecule type: DNA
 A-Residues: 1-382 <WIL>
 A-Cross-references: EMBL:268132; PIDN:CAA92221.1; GSPDB:GN00028; CESP:F09C8.1
 A-Experimental source: clone F09C8
 C-Gene:
 A-Gene: CESP_F09C8.1
 A-Map Position: X
 A-Introns: 7/1; 71/3; 109/1; 164/3; 192/3; 245/2; 263/3; 303/3; 360/3
 A-Superfamily: Caenorhabditis elegans hypothetical protein R07B7.8
 Query Match 7.9%; Score 61.3.5; DB 2; Length 382;
 Best Local Similarity 38.9%; Score 4.7e-34; DB 2; Pred. No. 1.9e-33;
 Matches 139; Conservative 60; Mismatches 111; Indels 47; Gaps 10;

Qy 376 SDTVPVSTHLRKPADINVIGALGDSLTAQ-----NGACSTPGNTLVDIYQYRGLSWS 427
 Db 37 SKKPYTPVNVNSVRPADIKLMLALGDSLTVGGFLFSSQAANGAGAE-DEPVAVLQYQYRGLAQ 94
 Qy 428 VGGDENITGTVTLANIREFNLSKGFCVGTCKETSNN--AFLNOAVGGRADLPVQ 483
 Db 95 AGGDTKTEERHVTPNLTQKNDTQVFGNGTC--SPNWTETARLNATAMPGANAKOPLQ 151
 Qy 484 ARRLVNDKNDTR-IHQBDWKITLFIGNDLCDPNDLYHYSQNFTDNIGALDILH 542
 Db 152 AROLVQLQQHETEVVNMEDWKUNLFIGNDLCDPNCQDQVQVILY 211
 Qy 543 AEYPRAFVNLTVEIYVNLRELYQEKKVYCPRMILSUCPCVLUKDDNNSTELATLIEFNK 602
 Db 212 DNVPVRVIVSLTGMLHEMLRQT-DTGHWFCQR-LHDECCG-----ESNK 254
 Qy 603 KCFP-----KTHQKLSGRVYRVEDTUVVQPFENNDMPKTSSEGULPNSFFAP 651
 Db 255 NFDADIRQACTDYNKFKQITDGTEKFDNDTQVQPMFOPTLPIVENGKPTQKFAP 314
 Qy 652 DCHFSSISHSRAASALWNMNLLEPVGOKTRHKFENK-NITCPNQVQPFERTYKS 707
 Db 315 DCFHFSQMGHALVSTYIWNILQPVGSKSTVSNSMSPVQTLACPDAACPFIRTPKS 371

RESULT 3
 T26083
 hypothetical protein W02B12.1 - *Caenorhabditis elegans*
 C-Species: *Caenorhabditis elegans*
 C-Date: 15-Oct-1999 #Sequence_revision 15-Oct-1999 #text_change 01-Mar-2002
 C-Accesion: T26083
 R. Swinburne, J., Ainscough, R.
 Submitted to the EMBL Data Library, October 1995
 A-Accesion: T26083
 A-Accession: Z220149
 A-Status: preliminary; translated from GB/EMBL/DDJB
 A-Molecule type: DNA
 A-Residues: 1-414 <WIL>
 A-Cross-references: EMBL:266521; PIDN:CAA91393.1; GSPDB:GN00020; CESP:W02B12.1
 A-Experimental source: clone W02B12
 C-Gene:
 A-Gene: CESP_W02B12.1
 A-Map position: 2
 A-Introns: 5/1; 61/3; 101/1; 156/3; 212/2; 273/2; 297/3; 332/3; 368/3
 C-Superfamily: *Caenorhabditis elegans* hypothetical protein R07B7.8
 Query Match 7.8%; Score 60.5%; DB 2; Length 414;
 Best Local Similarity 36.8%; Pred. No. 1.9e-33;
 Matches 141; Conservative 57; Mismatches 118; Indels 67; Gaps 10;

Qy 376 SPPVPPSVHLRKPADINIVGAGDSITAGNGGSTPGNVLVDLQTQGLNSVGGDENIG 435
Db 35 SKKPUPTSAHSVPADIKVGTGALGSLSLTANGAQAQGDPLAVLQTQGLAFQGDKSLD 94
Qy 436 TTTLANILREPNPSLKGPSVGTGKETS-PNAFLNQAVAGGRADLEPVQARLVDLKNID 494
Db 95 EHTITVANVLRKNPNTLUGASKGIGSENWENVSHLNMGVGAESKDITQARALVNTMH 154
Qy 495 TRTHFQEDWKITLTIFCGNDLDCFC---NDLVHYSQPNFTDNIKGALDTHAEVP--- 546
Db 155 SEINVKEWDKLVNIFIGANDICVYCEDPYFNSTALHGNATFKKNTIAVQIQLDNLRRT 214
Qy 547 ---RAVIN-----RATV-----RATV-----RATV-----RATV-----RATV 274
Db 215 NFKFKLFSRISVCKTFSSWREPKNSNYSRTTISLTGMNMRKIDKKKPCGEHTP 274
Qy 579 SLCPVCVLUKFDDDISTELATLIEKKFQBKTH-----LIESGRVDTREDFPTV 626
Db 275 B-CDC-----ENKQFDDQGCVCFYMNAAKDIQNTGLENDKNDPFTV 318
Qy 627 VQPFFENV-DMPKTSSEGUDPNNSFAPPCPHESFKSHSRAASALWNMLPEPYQKTRHKF 685
Db 319 VQPFFENGILDPPYASPQGVDMFFAPDCPHFHSAYGHGNIGMHLWNTIQPVQFKQTSVNL 378
Qy 686 EN-KINTCPNQVQPFRTYKNS 707
Db 379 SDPSVGLHCPSNCPEFPPTKNS 401

RESULT 4

hypothetical protein F13H8.5 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T16060
R;Ding, H.
submitted to the EMBL Data Library, July 1995
A;Description: The sequence of *C. elegans* cosmid F13H8.
A;Reference number: Z19830
A;Accession: T16060
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Residues: 1-548 <WIL>
A;Cross-references: EMBL:275955; PIDN:CA00110.1; GSPDB:GN00023; CESP:R07B7.8
A;Experimental source: clone R07B7
C;Genetics:
A;Gene: CESP:R07B7.8
A;Map position: 5
A;Introns: 4/1; 55/3; 91/1; 147/3; 223/2; 277/3; 335/3
C;Superfamily: *Caenorhabditis elegans* hypothetical protein R07B7.8

Query Match 5.2%; Score 406.5; DB 2; Length 348;
Best Local Similarity 32.2%; Pred. No. 5.9e-20;
Matches 117; Conservative 66; Mismatches 133; Indels 47; Gap 11;

Qy 1060 IKPAIENWIGSDFLCTEWKASNSV-----PTSVHQIARPDTKVAAALGDSLTTAVCARPN 1113
Db 15 VKMALLE-----KYNLSDIPHRIVYFDVNVMIKPHIRVIGAMGDSIT--IGSRAE 61
Qy 1114 NSSDLPTSWRLGSWSIYGGSMSLBTHTLTPNLK---KENYLIGGSTSTWEG-TAGLNN 1168
Db 62 NIVGQRQRYPGNAFFGMDFEVDRHLTIVNIFPRITAETKGKNUFGGSTGIDYGENGLN 121
Qy 1169 AAEGARARDMMAQAWDIVERMKNSPDINLEKMKVLTFIGND--LCHYCENPBAHLA 1225
Db 122 AIGGMKSDILRQAKSLVRSIKANAKSINTENDKUVLWIGTNDVGTLYQLEDP--V 179
Qy 1226 TEYVQHQIQLDILSEBLPRAFPVNVEVMEMLASLQOGGGKACMLAQANNCTLRSQOS 1285
Db 180 DEYKSHTEKGJLKLKENLPRTIVSIYGMFPQQLORAO-----SILKRGKA 226
Qy 1286 LEKOEKVKYNNLQSGISIISYWHQYTOR--EDFAWVOPP-FONTLTPNLERGDTDLT 1341
Db 227 RTVENQKLD-DLSGDSYRNVSDFQBNHENSNDFTWVOPFATEYTDYSYDHEGKYNPT 285
Qy 1342 FFSEDFPHFSORGHAEMAIALWNMMEPVGRKTTSNFTHSRALKCPSPSPYLYTLRN 1401
Db 286 FYASDFFLSKFGHAKHWTWLNLFEPVGKETKRDGTPKTYELNEKNLIKIVGN 345
Qy 1402 SRL 1404
Db 346 SKM 348

RESULT 6

T24015 hypothetical protein R07B7.9 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C;Accession: T24015
R;Harris, B.
Submitted to the EMBL Data Library, July 1996
A;Reference number: Z19830
A;Accession: T24015

Qy 653 CFHFSKSHSAAASALWNMMEPVGQKTRHKPENKI---TCPNQVOPPLRIVKNS 707
Db 326 CFHLSPTAHDIAKQIWKGLFEPIDOKT---ITNQLSVGFDRFVCPVBCPVLRTION 381
Qy 708 MOGHGTWLPCCD---RAPSALHPSVHALPADIOV 740
Db 382 ENCE---PSREMRFURVPSAFASSPGLSMPPMIPV 414

RESULT 5

T24016 hypothetical protein R07B7.8 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Accession: T24016
R;Harris, B.
Submitted to the EMBL Data Library, July 1996
A;Reference number: Z19830
A;Accession: T24016
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: DNA
A;Residues: 1-548 <WIL>
A;Cross-references: EMBL:275955; PIDN:CA00110.1; GSPDB:GN00023; CESP:R07B7.8
A;Experimental source: clone R07B7
C;Genetics:
A;Gene: CESP:R07B7.8
A;Map position: 5
A;Introns: 4/1; 55/3; 91/1; 147/3; 223/2; 277/3; 335/3
C;Superfamily: *Caenorhabditis elegans* hypothetical protein R07B7.8

Query Match 5.2%; Score 406.5; DB 2; Length 348;
Best Local Similarity 32.2%; Pred. No. 5.9e-20;
Matches 117; Conservative 66; Mismatches 133; Indels 47; Gap 11;

Qy 1060 IKPAIENWIGSDFLCTEWKASNSV-----PTSVHQIARPDTKVAAALGDSLTTAVCARPN 1113
Db 15 VKMALLE-----KYNLSDIPHRIVYFDVNVMIKPHIRVIGAMGDSIT--IGSRAE 61
Qy 1114 NSSDLPTSWRLGSWSIYGGSMSLBTHTLTPNLK---KENYLIGGSTSTWEG-TAGLNN 1168
Db 62 NIVGQRQRYPGNAFFGMDFEVDRHLTIVNIFPRITAETKGKNUFGGSTGIDYGENGLN 121
Qy 1169 AAEGARARDMMAQAWDIVERMKNSPDINLEKMKVLTFIGND--LCHYCENPBAHLA 1225
Db 122 AIGGMKSDILRQAKSLVRSIKANAKSINTENDKUVLWIGTNDVGTLYQLEDP--V 179
Qy 1226 TEYVQHQIQLDILSEBLPRAFPVNVEVMEMLASLQOGGGKACMLAQANNCTLRSQOS 1285
Db 180 DEYKSHTEKGJLKLKENLPRTIVSIYGMFPQQLORAO-----SILKRGKA 226
Qy 1286 LEKOEKVKYNNLQSGISIISYWHQYTOR--EDFAWVOPP-FONTLTPNLERGDTDLT 1341
Db 227 RTVENQKLD-DLSGDSYRNVSDFQBNHENSNDFTWVOPFATEYTDYSYDHEGKYNPT 285
Qy 1342 FFSEDFPHFSORGHAEMAIALWNMMEPVGQKTTSNFTHSRALKCPSPSPYLYTLRN 1401
Db 286 FYASDFFLSKFGHAKHWTWLNLFEPVGKETKRDGTPKTYELNEKNLIKIVGN 345
Qy 1402 SRL 1404
Db 346 SKM 348

Db 342 NTLISGVHDNITQIDSLSLGK-OTTNLINLEKOMVLRFVGQDLSPTPDK----- 388
 Qy 931 VLTRENSQELARLEAASRAVYSSMRVSGRSRYTDQDFSVLQLPQFRQNTLQPLVIALDG 990
 Db 389 -----TDFSTLETDNA-ESAWAKINTAS-FGKDNTISSTAYPKTINNLNSUVN 438
 Qy 991 PD-----TSPPAPDCIHQNQKFSQSLARALWTMPLIGSKYETLDRALAEMLPCTPNE 1045
 Db 439 GYDTIATLVNGSLTPQKOTYEAKUGVLAQYATAFGIKTRADSA---APTVN 493
 Qy 1046 PLRTPRNNTYPIKAIENRGSDFLCTEW---KASNSVPSPVHOLRPADIKVVAAG 1101
 Db 494 -----SNENFTKQITIVPK-----ATEYTIKLOASNGLNS-----GISIVG 531
 Db 1102 DLTIAVGRARNNSDLPSTMWSGDSGLET-HTTLWILKKENPVLGFSTSW 1160
 Qy 532 KTGTTGAIVNPNSVUTLDNRTRTS-----KGNTTATTSP-----MTVYUPLKVFGL 582
 Db 1161 EGSTAGLNVA--EGARAD-----MPAQA-----MDLVERMKNSPDINLEKDW 1201
 Db 583 QASGTVTASVONGSGGDSVNPFLIPANASSEYAGGSQSYISTLN---NIDTA 638
 Qy 1202 KVLTFLGIVNDICHYCENPEAHATEYVHQIA-----ADI-----LSELPLPAFV 1248
 Db 639 SJLTFLYGAPG-ATYSOKSCKTKVSDFEASQOSIFKWMGNMDVTQIMSRUSSDDVQFM 697
 Qy 1249 NVVEVMELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1308
 Db 698 NY-----GKANIKNTDPLTLLNNTIAIQLKEA----TLSNLPS-NYF 737
 Qy 1309 HQYTOREDPAVVQOPFFONTLPLN 1333
 Db 738 SQVSSD-----LQAWINQTMOTLN 756

RESULT 9

T38084 TRAP-like protein - fission yeast (*Schizosaccharomyces pombe*)
 C;Species: *Schizosaccharomyces pombe*
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999
 R;Gentile, S.; Odell, C.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
 A;Reference number: 221768
 A;Accession: TJ084
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-355 <GEN>
 A;Cross-references: EMBL:268136; PRION:CA922391; GSPDB:GN00066; SPDB:SPAC1F5.11c
 A;Experimental source: strain 972b; cosmid c1F5
 C;Genetics:
 A;Gene: SPB:SPAC1F5.11c
 A;Map position: 1
 A;Introns: 22/1

Query Match 1.9% Score 149; DB 2; Length 3655;
 Best Local Similarity 20.0%; Pedi. No. 1.5; Mismatches 307; Indels 248; Gaps 40;
 Matches 167; Conservative 112; MisMatches 307; Indels 248; Gaps 40;

Db 509 MYGL-----TCRKRCIETGGEOFSGFLTNIKPFEAVTF-----OKLFFEVGK 551
 Qy 895 VLVNLVUDPEFLNPTIMQFELGNDKCPVQQAISLNCVLTRENSQELARLEAASRAVYSS 954
 Db 552 GF-----SFRRP-----EQYLETFC-----EESLWRPAISTPLPNKDEKUCLEVATIF----- 599
 Qy 955 MRLVUGSCRYDQDFSVLQLPQFRQNTLQPLVIALDGPD-TSPPAPDCIHQNQKFSQSL 1008
 Db 600 -----THLEPSIFLKOFETN-LPTFPDQLKGULFHIPQELLSNSTSSKFL 646
 Qy 1009 SOLARALWTMPLIGSKYETLDRALAEMLPCTPNE-----TCPTONEPF----- 1049
 Db 647 NILLRFL-LSRIBELGSS---DIRHGSVLLRFLRSFTVMSNFATENEPVLPVYSETI 701
 Qy 1050 -----TPRNNTYPIKAIENRGSDFLCTEW-----KGNTTATTSP-----MTVYUPLKVFGL 1100
 Db 702 VKCMKLMANSANSLNVYTLLRFLRGIGGGRSLEYK-----EMPLMLAALLEPFLNL----- 755
 Qy 1101 GDSLTTAVGRARNNSDLPSTMWSGDSGLET-----WRLGSWSTGGDNLET----- 1137
 Db 809 TPEBLCDLWLTDPDFRIMAPYIBDLMLNWLQPLPYNHSHALKLKGKGRK 868
 Qy 756 -----ISARTPKEKDFLTELCLTFLVRLSLLPPLWMSYLPLVMSIKSSOBLVSGLR 808
 Db 869 -----LDRYQSLSPEN-----NOFLTSLIKGKQPOKL-----KTYQYBAV 910
 Qy 1138 -----HTTLEPLKLF-----NPVYLGESTSW-----EGTAGLNVAEGARADDM 1178
 Db 809 TPEBLCDLWLTDPDFRIMAPYIBDLMLNWLQPLPYNHSHALKLKGKGRK 868
 Qy 1179 PROALDUVERMKNSPDINLEKOMVLUFLCLTFLVRLSLLPPLWMSYLPLVMSIKSSOBLVSGLR 1236
 Db 911 NLJS-----PSDLEVKQOQAFYVONISKLYVVKYKSDATNSLASSIRCTADKISKSNPDRPR 968
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1292 KRNWNLQGTSFSYPMYQTOREDFVWVOP-----FQONTLPLMFRGDDPLTSE- 1345
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC 1069
 Qy 1237 DLSEELERPLAFYVNUV-----VMAELASLJYQGQGKCAMLAQNCCTCLHSQSSEKQLKVNWLQHGSSSYW 1291
 Db 969 YSVIPSRMTGGRSFT--QLSDDSDTEITLASATYGLFATWDELNB---BAYFWIEK 1021
 Qy 1346 -----DCPFRSD-----RGHAEMAIIWNMLPEPVGRKTTSNNFTSRAKUJC 1388
 Db 1022 LAVNVTVHDFIAFDVQGNSKFTTNIKEV-----ISPHYFACILSEVVC

A; Reference number: A59328
 A; Contents: annotation
 A; Genetics: XFI952

Query Match 1.9% Score 146.5; DB 2; Length 1755;
 Best Local Similarity 18.7%; Pred. No. 0.63; Mismatches 457; Indels 465; Gaps 55;
 Matches 245; Conservative 142; Mismatches 457; Indels 465; Gaps 55;

QY 193 VDSLGMVLDYLQQEVPRAFVNVLSEV--AEVSRQHGTW-LSPAPCPCNCSEETRL 248
 Db 57 LDRALG--OARBEIEFENPVPTDSRILSCADYLHQVGTLMELYAPAWYAEEMRL 113

QY 249 AKVVMQWSYQEAMNSSLASSRYSQESFTVVFQ---PFFYETPSLASEDPRLQDST 303
 Db 114 AKWQ-----VDGVNDRREACAVLMLRGCTLLPDYLRLQNLONGHRDIP----I 155

QY 304 LAMHILWNRMEEPAGEKDEPLSVKGRPRMCPSPRSVPFSYRNSYLRLQKQDKLVR 363
 Db 156 VLIPLNLNEIRAAARGE-----SRMENVLFAFDP----EVR 186

QY 364 EGAEIRCPDKDPS-----DTVPTSH---RLKA-----DINV 394
 Db 187 NATEVDELHARSSLTGCNRQLDPAASATKEELVRKALDLVLRIGSDVTUROTEL 246

QY 395 GALGDSL-T-AGNG-AGSTPGNVLDLTVQYRGLSMVSGDENI--GTVTTLANTIREFPNS 450
 Db 247 SSVADTLAMGLMGSL-----VLOQRDALLRIVESDEKABALLRIGAHLVDS 300

QY 451 LKPSRVGTCKETESPNAFLNQAVAGGRASDLPVORRLYDLMKNDTRIFQEDAKITIIFI 510
 Db 301 LDNOVANILGABAADADVLANA-----ETRRTLUDVLAIBAISNTAACFCIAFI 349

QY 511 GGNGLDCFDCLDLYHSQPNFTDNGKALDILHAEVPRAFVNVLTVLEVNRLRYQEKV 570
 Db 350 ESN-----WDQARLKVPGILSVEYGGASILELSQSASTYLEV-----RUYI 391

QY 571 YCPMRMLRSICPCVLFKFDNNSTELATLIEFNKKQEKTHQOLIESGRYDTRDFTVVVOPF 630
 Db 392 EC-BLITERPVPAARQDPLADAMASLEYLEAIRE-----RRCQDILEITNS 441

QY 631 FENYD--MPKTSRGLPDNSFAPDCFHSSKSHSRAASALNNMLPVGQKUTRHKEN 687
 Db 442 LEALSYWPIP-AEGDSVGSQSQSTDVLTFPSAPD----BALPELTUVVEDGTVTFOEAWSQ 495

QY 688 KINITCPNQVQPEURTYSKMSMQHGTWL-----PCRDRASALHTS 729
 Db 496 SLTIVEPESADIEPASLRLVQPTVTDLXEGDRRSELPESTMPDRPDEPVQSIPTVVAUL 555

QY 730 -----VHALRADIQVVAALGSLTAGNG-----IGSKDDLPDVTOYRGLSYS 774
 Db 556 SGGRFDDPGVVIDOSDIPFL--EELDEFOGNURRLTWTGTAFFDM-DSLRQIRRFT 612

QY 775 AGGGSLLENVTLN-----ILREFRNRLGY 801

Db 613 LKGSGRLVGAALLTGHPAWKTEINLINRVLDGSSRAASPAVVALVERACEVVLPENNAALKK 672

QY 802 AVGTCGANDTNAFLNQAVGKAEDLMSQVQTLMKMOKMDKHDHRNFEDKIVTLLIGSD 861
 Db 673 ASISTDBATQALIDLVASGEA-----FTYSPSDIASDAT-----LKLVPIQLEG-- 719

QY 862 LCDYCTDSNLYSAANFVDHURNALDVHLRREVPRVLUVLDF-----NPTIMQV 911
 Db 720 -----GABPFVDSLRL-EILEABIANHLDITKDWINTAGQSYNSVELRAV 766

QY 912 FLGN-----PKCPVPOO-----ASVLCMCVLTRE 936

Db 767 HTLNGAFAMTEMPEITQVOMOSAETYIKRJLTHKHOKASVEGVGVLSAMATAIDITALRS 826

QY 937 NSQSLARLEAFSRAVYRSSRELVSGSGRYVQEDSVLQPFQONI--QLPVLAQDGLPT 993
 Db 827 DS---PRIFSPV-SLVSCISELV---TWPMDQGOSQOQYFEEVQAONLKSYPEQTODV 877

QY 994 SFFAPDCIHPNQKFSQLARALWNMIEPLIGSKTETLDRAMPICTQNPFLRPPRN 1053
 Db 878 SL-----SNQVLTDPVYDLSNLDVLPN-----PERPIVNRQOE----- 917

QY 1054 SNTYTPIKPAIENWSDFCLTEWKASNSVPTSVQHQLRADIKVVAALGDSLTTAVGARN 1113
 Db 918 -----VHQY-----DNVAAST----- 929

QY 1114 NSSDLPTSWRGLSWSICDGDNLETHHTLIPNLTLLKFNPVLLGFSTSTWBGTAGLNVAABGA 1173
 Db 930 -----MSDVVAADTNL----- 940

QY 1174 RAADMPOAQADUVERMKNQSPDNL-----EKDKWVYLFT-GVNDLICHYCENPEAHAT 1226
 Db 941 --DVAHQDYDVLVSG-NESPANEBLGFLAFDERAPFLVLDLFBESSDLDLHDCONLAKU-- 994

QY 1227 EYQHQIQLDILSBELBRAFPNVVNEMLASLYQOGGKCAMLAQNCNCLRHSSO- 1285
 Db 995 -----NEAQDRELLVQGR--DLHTLGK--ARMAGNAIGDGHSTBSM 1036

QY 1286 -----LEKQELKVNWNLQHGISSFSVWQ-YTQREDPAVWVQ 1323
 Db 1037 LESVMADYIMLNRRDMRILYE-----SFDV1QHMLQTROHRVWTP 1078

RESULT 11

T5278 Hypothetical protein R155.4 - *Caenorhabditis elegans*
 CSpecies: *Caenorhabditis elegans*
 CAccession: T15278
 CDate: 20-Sep-1999 #sequence_change 20-Sep-1999
 A;Cross-references: EMBL:AF003390; NID:92088866; PID:92088867; PIDN:AAB54270.1; GSPDB:GN
 A;Experimental source: Strain Bristol N2; clone R155
 A;Reference number: Z28321
 A;Gene: CESP:R155
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-913 <GET>
 A;Cross-references: EMBL:AF003390; NID:92088866; PID:92088867; PIDN:AAB54270.1; GSPDB:GN
 A;Genetics:
 A;Accession: T15278
 A;Intron Position: 3
 A;Map Position: 47/1; 68/1; 355/1; 719/1; 794/2; 864/1

Query Match 1.8% Score 141.5; DB 2; Length 913;
 Best Local Similarity 19.2%; Pred. No. 0.44; Mismatches 335; Indels 339; Gaps 48;
 Matches 191; Conservative 132; Mismatches 335; Indels 339; Gaps 48;

QY 646 NSFAPDPDCHFSSKSHSRAASAL-----WNN-----MLEPVGQKTRHKF----- 686
 Db 2 SEFF-----FLAMANNHSNFAESLSLTIWLNICHAGSTPLVNPMSGIQLQEPHDFNFGI 57

QY 687 --NKINTCQVQPEURTYSKMSMQHGTWLPCDRDRASALHTSVHLPADIQVWAL 744
 Db 58 GLNLLARRT"KDELSTFICQPSRSM-----LLT 85

QY 745 GDSITAGNIGIGSKDD-----LDPVTVQYRGLSYSAGGDG- 779
 Db 86 GISQGQGLAAGSIPPDDELJSEFUNGITTTQARIIPVVKLTEMVQKNTLEPVSNDL 145

QY 780 -SLRN-----VTTLPNLBEPFRNLUTGYVGTGP--ANNTNAFLNQAVP-----GAKED 826
 Db 146 KTENHYLAMSSITADKVKGFK--GAYVVPDGDPYTNLDFKVSQYTLDIESLAENLK 202

QY 827 LMSQVOTU-----MQKMKDHRN----- 845

Db 203 PISSEMUKTDITKTTHKMFYKIOSAITTISDIKVMDFQSGDSALNVTKTSLPELSS 262

QY 846 ---FHEMKVITVLIGGSDLCDYCTDSNLYSAANFVDHRL-----NALDVHRE- 891
 Db 263 VVSAVSEHEKIGKRTISDSNLAEP--SKSMSSSNASBELNKHATDLCKLNQMFVYHFK 320

QY 892 ---VPRVLVNLVDFLNPTIMRQFLGHPDKCPVQASVLCNCVLTRENSOBLEARLF 947
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 321 HGNKVKPVKPIGLD-GHDTTIFAEN---SWLNDAVDELCSCEIGLSGV- 368

QY 948 SRAYSSMRELYVSGRYTO-BEFSWVHQPEFONIQ----LPVLADGLPPDSFEADC 1000
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 369 --ATDQAKKEIRKIGSLAKLQLSALISPLASKVQD1KIGSLPAESDK-- 415

QY 1001 IHNQNKFHSQQLRALWTNMLEPLGSKSTETLDLRAEMPTCPIONEPFLTPRNNTY-- 1058
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 416 ----NLTEKALITON-ETLTSKEYALN--EM--VKTIND-LKKONSDALORM 459

QY 1118 LPSSWRGJSWSIGGDNIAHTHTTLPNTLKENYLGEFTSTWEGTAGLNAVAE----- 1171
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 504 QPMTM-----KTTAESTADALK--NPDVLATANQTPEVINSNTPEIKPVIS 548

QY 1172 --GARARDMPQAQDWLVERMKNSPDINLEKOKWKLVLFIGNDLCHYCEPEAHATE 1227
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 549 ATKAFPEETSVIY--ITSVKILPD--KOMKNLNTPEVKNN-----SNKTK---- 593

QY 1228 YWQHIQQALDISBELLPRAFPVNVEMLASLYQOGQGKSCAMAAQONCTCIRHSQSSE 1287
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 594 --ESSVDL-KELKVA--TVOLSTIGSVARGFRMEQAGLTLNDVSNMK-SFEAVV 643

QY 1288 KQELKKVWNLRORGISSYSWYHOTREDFAVTV-----QPFQNTLPINER 1335
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 644 KQEMSKVKLD-----GTDKQMKMVKLGLGDBELEKULSDNLFVSKVPNSN 690

QY 1336 GDTDLTFFSEDCCHFSDRGAEMA-----ALANMMAPVGKRIT 1375
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 691 NLAQDFAGIFEAALKVKGYPHEVIAKSVLKMEAANSNSLAGLIPTINN 749

QY 1376 SNNFTHSRAKLCPSPESPYI-YTLRSRLR_PDOAERAPEVU--YMAVPUVAGVGLVVG 1431
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 750 SSFSFSKSKOSKLUDTFASFLNNTINNSRLASKTNAQSEPGALAKHTIVIVAGIPLII 809

QY 1432 I---IGTVWRCRGRG-----RDP-PMSLT 1455
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 810 IAVCVGVWVWRRRSQRRAEAFAHRNDPPTPSRRS 846

RESULT 12

F69189 protoporphyrin IX magnesium chelatase (EC 4.99.1.-) - Methanobacterium thermoautotrophicum
 C;Species: Methanobacterium thermoautotrophicum
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Nov-1999
 C;Accession: F69189
 R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, J.; Dubois, J.; Aldredge, T.;
 Oiu, D.; Spadaro, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
 A;Reference number: A69000; MUID:98037514; PMID:937463
 A;Accession: F69189
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-1708 <WTH>
 A;Cross-references: GB:AE000847; GB:AE000666; NID:92621756; PID:AB85178.1; PID:9262175
 C;Experimental source: strain Delta H
 C;Genetics:
 C;Gene: MTB73
 C;Superfamily: Methanobacterium thermoautotrophicum cobalamin biosynthesis protein N
 C;Keywords: lyase

Query Match 1.7%; Score 131.5; DB 2; Length 1361;
 Best Local Similarity 19.7%; Pred. No. 4.3; Mismatches 498; Indels 379; Gaps 58;
 Matches 249; Conservative 140; Mismatches 498; Indels 379; Gaps 58;

QY 593 ELATLIEENKKFOEKT-HOLIESGRVDTREDFVVQOPPFENVDMPKTSEGIPDNEFFA 650
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 181 DLTSLIAYKQKPGNTAHQWDCALEYAAGKTNLQFKALMVTNMRGIPWNSSWE 240

QY 651 PDCPHFSSKSHSRAASALTWNMELPVGQKTRHKFENKINTCPNOVQVQPFLR----- 702
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 241 P-----ATWPASPISSEFFYRDGQRTKFDYFTRYPLDPAKPTAVLSTVGSTCEV 292

QY 703 TYKNSMCHGTWILPCRDAPSALHPTSHALPADIQVAAGDSLTAQNGGGSKPDILP 762
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 293 TYADAMQ-----QIDALVSRGLNVNIVPGI--TWSNY----- 323

QY 763 DVTOYRGSLSYAGGDGSLENVTLPNLREF-----NRNLTGAVGTDGANDTNAFM- 816
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 324 -----ILNQGQMNTIOTLCPMPNQNTNITAIR-GIGNYDLSITGV 364

QY 817 QAVPGAKAEDLMSQVOTIMQKOMKDHRVNFHEDWKVUTVULGSDLCDYCTPSNLYSAN 876
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 365 TSVSSAN-----VIVEQIL-----DNGNLNIKSISTV-----QPNVISA-- 400

QY 877 FVQHLRNALDVHREVRPLVNLVDFLNPITRQVFLGNDPKCPVQOVSVCNCVLTURE 936
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 401 LUKPLTDASVVOYE-----ANPEKYPV-KANVILDMTEITG 437

QY 937 NSOELARI-EASRAYSMSRMELVGSRIFT-----QEDFSWV-----LOPPFQNIQ 982
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 438 STTSGASVTKFEDRSNPEPLRAMITSSYRTIQWVSEEGSMSVWQCAPEMOC-Q 496

QY 983 LPVLADLGHDPTSFAPDCTPHQNKFHSQLARALWTNMLEPLGSKTETLDLRAEMPTCPT 1042
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 497 IEPLAIGVGE-----ISGDPET-GAODITVITE 525

QY 103 QNEPP-----LRTPRNSN-----YTPPIKAENWQSDFLCTEWKASNSVPSV 1086
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 526 RIEKLVSAFNFNWLQTMWNSDKRVAIVYNNP--PSKONIGASYL-----NVPSI 575

QY 1087 HOIPLADTKVAAALGDSITTAVGARNPNSDIPTSW--RGLMSIGGDNLEHTTCEN- 1143
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 576 IEI---LKRKMGEGYS---VGEIPQDADALVEMMKGNGINVANWAGELEKANSNA 627

QY 1144 ILKCKFNPYPLGFR 1157
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 628 ILNPYEDYLAWFNT 641

RESULT 13

T30884 neural specific DNA binding protein - African clawed frog
 C;Species: Xenopus laevis (African clawed frog)
 C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999
 C;Accession: T30884
 R;Marine, J.C.; Bellefroid, E.J.; Samain, H.P.; Martial, J.A.; Pieler, T.
 Submitted to the EMBL Data Library, December 1995
 A;Accession number: Z20918
 A;Accession: T30884
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-1361 <MAR>
 A;Cross-references: EMBL:U42462; NID:91150837; PID:91150838; PID:AAA98467.1
 A;Status: preliminary; nucleic acid sequence not shown
 A;Molecule type: DNA
 A;Residues: 1-1708 <WTH>
 A;Cross-references: GB:AE000847; GB:AE000666; NID:92621756; PID:AB85178.1; PID:9262175
 C;Experimental source: strain Delta H
 C;Genetics:
 C;Gene: MTB73
 C;Superfamily: Methanobacterium thermoautotrophicum cobalamin biosynthesis protein N
 C;Keywords: lyase

Query Match 1.7%; Score 131.5; DB 2; Length 1361;
 Best Local Similarity 19.7%; Pred. No. 4.3; Mismatches 498; Indels 379; Gaps 58;
 Matches 249; Conservative 140; Mismatches 498; Indels 379; Gaps 58;

QY 56 NMPSKSVS---LKPDIKF---VAAIGNLBIPPPRTGDLKQDMTERQQC--- 103
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 10 NSPLTLMWSAAGLSPAEVHLKENGIGYIAPPGPPT-----TPTEYCHQMA 59

QY 104 -----MGWMTVLS-----DIRYFSPSPVMPVCHTGKRVIPHGDAE 139
 C; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;
 Db 60 PLASHASPYGYELVQSAAGNTSHLHDYLTMDVSRSSPRVTPLR--SRKRALISIPLS 117

QY 140 DLWTOAQELVNRMKENLQLDFQFDWKLINNFFNSNASQCYLCPSAQONGLAAGGV----- 194
 Db 118 DASIDLQTMWIRTSFNSL-----VAYINNSRSSAASGSYGHLSGAISSPASF 165
 QY 195 -----ELMGVLVDLQLOQEYVPRAFVNLDLSEVAVSROHG--TWLSPAPEPNCSETT 246
 Db 166 PIHKPCSLSAALSO-ORSLSSFEGHTPLHSPPTFASRQGALTSANPAPPNNSSAPDS 224
 QY 247 RLAKVUMQMSYQEAWSLL--ASSRYSEQEFTVF-QP----- 282
 Db 225 VLNKVSSEASAVSSTVNQVHRSKVKTEHADSPFPPDHLTDKEDLDKDBCKOEP 284
 QY 283 FYFVETTPSLHSEDPRLQDSTTLAHLWNRMMEPAGEKDE-----PLSVKH---- 327
 Db 285 HIVETINCHWDGCSKEFDTQDLVHINNDHH-GEKKEFVCRWQDCSRBQPKPAQML 342
 QY 328 -----GRPMCCPSO----- 353
 Db 343 VHMRRHRTGEGKPHKCTFEGCFKAVSRLENLKKTHLRSHTGEKPYCDHEGCNKAFAFSASDR 402
 QY 354 QKPODKLVEREGABI-RCDPKDPSDTIVPTSV-----HLRKPADINVIGALGD 400
 Db 403 AKHQNRTHENEKPYICKVGCTKRYTDSSLRKHKVTKVHGPEARTVTKHNRDIIQK-PS 460
 QY 401 LTAGNGAGSTPGNVLQDVLTYQRGLSWSVNGDDE---NIGTVTIANLIREFNLSLKGSV 456
 Db 461 LPKENG-----DNEASAKISGRERSDVSRSRQEBCLQPTTICBEDNMHQSP--GGOS 512
 QY 457 GTGKETSPNAFLNQAVAG-----GRAEPLPVQARRLVLMKN-----D 494
 Db 513 SCSESEPSPYGNNTNNIDSGVDSLAWQGS1GDLRFGLBETSPPVSDTSSWORSGRPATP 572
 QY 495 TRHFQE-----DKIILTFIGGNDLCPFDNLVHYSFONFTNDIGKALDILHA 544
 Db 573 ORHSAETGTAEBERIKONFERFLIYERPATC-----NTRLPTIANGFVIGVSSVNLN- 629
 QY 545 VPRAF--VNVLVTELEVNRLVELYQBSKKVYCPRMILRSLCPVCLFKFDNSTELATLFN 601
 Db 630 -PRAIELSANDVTMNMQANERR----- 657
 QY 602 KKRQEKTQHOLIESGRYDRTFEDFTVVWQPFENNDMPKTE- -GLPDNSFFAPCPHFRSK 659
 Db 658 -----SSAYTSRR--SSGISPYFSSRSSETSQFGGRINNNSASDYP1STD 703
 QY 660 SHSRAASALWNMILEPGOKTRHKFENKINTC-----PNQVOPFERTYKSMQ 709
 Db 704 : ASRSESEAQSQHSLPNLMLTIPACHYRILAKYAATGGPPPTPLPNMRIGLANKLSD 763
 QY 710 GHGTWLP-----CRDRAPSA-LHPTSHALRPADIQVVAALGDSLTAQNGISKP 758
 Db 764 GADFLPLPPFQLPVPRRCSDGGMAGLIPMWHRI-PGNSRRAA- -DPVRTRAGIDXP 820
 QY 759 DLDEPDVTTQYRGISYSAGCGDSLBNVTTL--PNLTREERNLTCYAVNGDAN----- 809
 Db 821 -LPRPSRFH-----SMNSMNTLHPSLSERRNGLQHYTCSDGGHLRHVSPR 867
 Qy 810 ---DTNAFL---NQAVPGAKEDIM---SQVOTLQMKMDKDHMFHEOKVITLI 857
 Db 868 PRSTSENVAMEATSCDADYPGGD-DDLMPLDDVWYOTINSQRAPEQMLTEVS----- 920
 QY 858 GGS DLCDYCTDSNLSAA---NFVDHLRNALDVLHREVRV-----LVNLDFNAPT 906
 Db 921 -----SPARNUQSNTKSFHNNTPEQPRAGVLSRNFPALAECLGQTAQMDNMPV 972
 QY 907 IMRQVFLGNDP---KCPVOQOASVLCNCVLTRENSOELARLEAFSRAVYSSMRELVGSCY 964
 QY 973 QWNEVSSGTVDSLPLKOOPA--GNLAVVQOKQNFQAQYQSNQAMQRANIMGOCB 1029
 Qy 965 DTQEDPSVLUQPF-----FONIQUPVIAQGLDTSFFAFCI--HPNQKFHSOLA 1012
 Db 1030 SVORNISTNGQRNYLQOROOQOMSOCOI-VSISIPIOPRYSOSOMSUSRAMOBGOSIS 1088

QY 1013 RALWTNMMBPLGSKT---ETLDLRAEMPTCPTO-----NEPLRTPRNSTMVYTP 1059
 Db 1089 PSC-NNMTPRPGVTHAIPSNTLNHQRLAVHGAFTQGPANNFSVNQDGLHPP-NAYTVQP 1146
 QY 1060 IKPAIE 1065
 Db 1147 QKNGLE 1152

RESULT 14

E96590 hypothetical protein T24C10.3 [imported] - *Arabidopsis thaliana*
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)
 C;Date: 02-Mar-2001 #Sequence revision 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: E96590
 R;Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Ainsen, N.F.; Hughes, J.; Huijzer, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marzali, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shiota, T.; Shiu, P.; Southwick, A.M.; Sun, H.; Tallon, I.; A.; Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.; A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A;Reference number: A86141; MUID:21016719; PMID:1130712
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-891 <STOP>
 A;Cross-references: GB:AE005173; NID:9857517; PIDN:AAG00872.1; GSPPDB:GN00141
 A;Gene: T24C10.3
 A;Map position: 1

Query Match

| Best Local Similarity | Score | DB | Length |
|-----------------------|-------|----|--------|
| 19.4% | 130.5 | 2 | 891 |
| Conservative | | | |
| Matches 133; | | | |
| Mismatches 92; | | | |
| Indels 290; | | | |
| Gaps 171; | | | |
| Db 27; | | | |

QY 156 LQDFQFWKLNIFVLFNSNASQCYLCPSAQONGLAAGVDEL-----MGVL-DY 202
 Db 159 LNQGTERERVKASVNVNLSNTIEVTLDPNNSDLCPADSKDDDDQDOWEFKTAESMFGTUGGSY 228

QY 203 LOQEVPRAFVNVLDSLSEVAEVSRQYHGTWLSSAPEPCNCSEBTRL--AKVVMQ----- 254
 Db 229 KOBEREKVONTADSS-----GWSSPAINGCPRNPETAKUDVAKVLAERENGDD 279

QY 235 --WSYQEAWNLLASSRYSEQE-----SFTVUQPF--YETTSLHS---EDPRL 298
 Db 280 DPMD-NGGWEFKVPAEKEPKRDLINKKGNGWGFGFPEVSKLETTNSFOSSVEKTKM 338

QY 249 Q-----DSTIATWHWNRMHPEPAGEKDEPLVKHGRPMKCPBPSQSBSPYLSYRN 346
 Db 339 ENGSISFPNSGNDVNSGGTSWAFKOPSLIGNEKEK-EVOTSK---PKGVPLSFEDDE 393

QY 347 SNVLTROLKPKDQKLVREGAEIIRCPDKPSDWTPTS--VHRL--KPADINVIGALGSLT 402
 Db 394 KSTSDTIVHEDSFVLSDFPVERKTKAPSPTVISDLSRISYQSEKNAILESAT 453

QY 403 AGNGAGSTPGNVLQDVLTYQRGLSWSVGDENITGTVTIANLREFNSSLKGFSVGTQET 462
 Db 454 ASNEVNGEDD-----SWEFOGPKMFTDSGJIAEGADD-----STWEFO 493

QY 463 SPMALFLQNAVAGGRAEIDLQVQARLVDLQMDNTRHQ--EDWKJLTFGNDLDFC 519
 Db 494 GRPAALKNS-----DYEVNUBFDDDS-WEOQGTOPVKOSMRICDGLWYR----- 540

QY 520 NDLVHYSQNFNTNIGKALDILHAEVPAFVNL---VTVLEVNBLBELYQKVKVYCPM 575
 Db 541 ---KHSSTENEVGN-----QSSVNGFGSLHDKVTRIEPNDYKDLFHKKLIEYH 589

QY 56 TIRSL-----CPCVLFKFDNSTELATLIERNKKEFOBKTHOLIESGRYDTHDF 623

Db 590 ALYTHELEKKEARUAKAASDEVQKCDSEFEDIQNLNNNDVLISGVNLLSQPESSGMELY 649
 Qy 624 TVVVQPFENN-----DVKTE-----GLPPNSFRAPDC 653
 Db 650 KALQEPRPRELSEDLLTERLISAEKWSTTBLKHATLTUKINLGSLEQSKYASTW 709
 Qy 654 FIFSS--KSHSAAASALWNMMLPVGQCTTRHFEKNINTCPNQVOPFLRYKNSMOG 711
 Db 710 FETTSCTAOELRHAASTIKVNDKVEEILSKPQKSVALSVGEIYRVVKILASTR 769
 Qy 712 GTWLPCDRAPSAHPVSHALRPAD 737
 Db 770 KWLI-----LLAPISSNVLAVID 787

RESULT 15

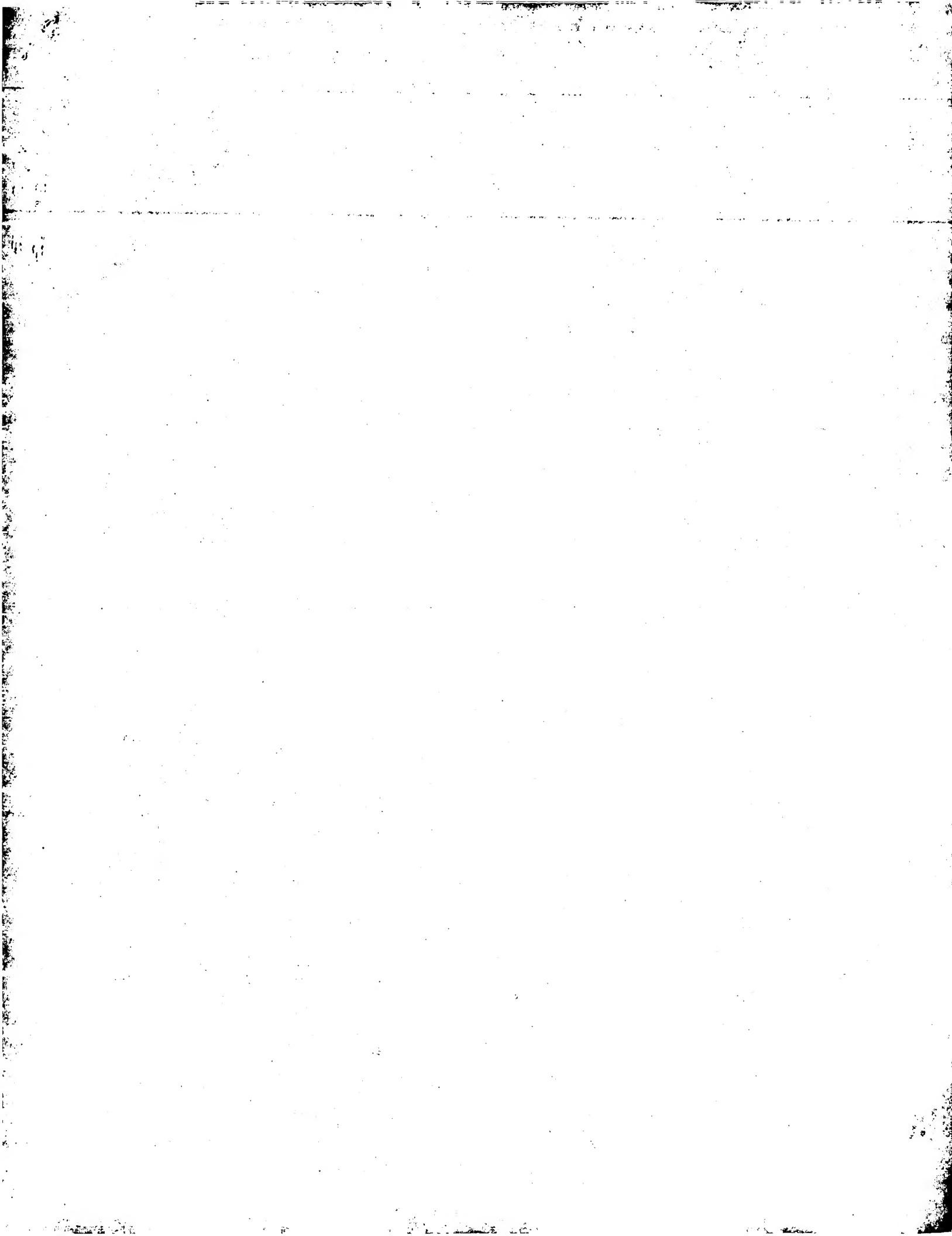
T15279 Hypothetical protein R155_3 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C;Accession: T15279
 R;Geisel, C.; Wamsley, P.; Kramer, J.
 Submitted to the EMBL Data Library, May 1997
 A;Description: The sequence of C. elegans cosmid R155.
 A;Reference number: Z18321
 A;Accession: T15279
 A;Status: preliminary; translated from GB/EMBL/DDJB
 A;Molecule type: DNA
 A;Residues: 1-165 <GB>
 A;Cross-references: EMBL:AF003390; NID:g208866; PID:g2088870; PMID:AAB54273.1; GSPDB:GN
 A;Experimental source: strain Bristol N2; clone R155
 A;Gene: CEPB;R155.3
 A;Map position: 3
 A;Introns: 662/2; 741/1; 813/2; 941/2; 998/1; 1049/3

Query Match 1.7%; Score 130.5; DB 2; length 1165;
 Best Local Similarity 19.5%; Pred. No. 38; Mismatches 185; Conservatism 140; Mismatches 356; Indels 267; Gaps 44; Matches 185;

Qy 396 ALGDSLTTAGNGAGSTPGNV--LDVLTQYRGIS--WSVGGDENIGTVTLANILREFNPS 450
 Db 15 AVKDELNLTKLGAVDMWSINTLIDELSNFLESSRLSMNDEERHLHLIKLKAASVYTQTL 74
 Qy 451 LKGPSVSGKETSPN----AFLNOAQAVAGGRA--EDLPVQAR--RLYDLMKDTRHFQ 501
 Db 75 LKVNINYSQETKVNSEVVYSSLQKVYTSIASELPSKASNGDKIDMNS----FEK 129
 Qy 502 DKMLIT-----LPIGGMUDCFCNDL---VHISP 527
 Db 130 SIKSLTEDSPSITGTYEULSKSRHLRVSSVALKHVGFLNGHSIDLHFKDQEDKWKDAV 189
 Qy 528 QWFTDNIGKALDLIHA----EVPRAFVNLYTVLEVNRLRELYQEKVYCPERMILRSLC 581
 Db 190 RHSGNUNALVQLKTEASTIKVHQFSVSEAVMLEVSK----IRSKI 234
 Qy 582 PCVLFEDDNSTHATLIEFNKPKOEHOLI--ESGRYDTRDDFTVUVQPFENNVD--MP 637
 Db 235 P---KYSGLADKIATL-----PSKTHEILGTPSVKLPTSCK-----DKVKAIP 275
 Qy 638 KTSSEGLDNSPFPAPDCPHFSSGSHRSAAASALWNMMLPVGQCTTRHFEKNINTCPNQV 697
 Db 276 LMAQ-LPAQKLAKLKLUMPSBQKLPENAKQGNIITIVNO-----DGK-----NOL 321
 Qy 698 QPLRLTYKNS--MQGHTWLPCDR---ADPSALHPTVSHALRPAD-IQVVA-LGDSL 748
 Db 322 ENVILOMKSSTVKEVYLLVLPKDTBLKONPIHTAANTISSLDKVERAQMADFF 381
 Qy 749 TAGNGICSSKPDG--LPDVTVYRGLSYSSAGCGSLENVTPLNPREFNRMLTGAVGT 805
 Db 382 TALNSIROEESLILPVITI-IQQLSLNDQFGLTIVNNLPMIKTOLS-SLTSY---- 435
 Qy 806 GANDTNAFLNQAVPGAACAEDMSQVQTMQMKDDHRVNFIEDWMKVITVJGGSDICDY 865

Db 436 -----INSVKKTKHAE-----TDALKDFEKGILS--RVIGTATRGI---- 470
 Qy 866 CTDSNLYSSAANVUDHRLNALOVLHREVPVVLVPLNPTIMQVFLGNPDKCPVQQAS 925
 Db 471 ---SNMOKLUDFKD-LADIGOLVKSEVKEVKVQDNL----- 502
 Qy 926 VLNCNCVLTRENSQELARLEAFSRAVRSMRRELVGSCRYDIOEDFSVVLQPFQNTQLPV 985
 Db 503 -----ENVANLKGALAGIEGQKLTASGBIVG-----YKKSVPV 535
 Qy 986 LADGLPPTSFRAFDCHP--NOKEFISQALARLWTNMLERPLGSKTERLDRAREMPITCPT 1042
 Db 516 TTSLSLPDYSTFFLDAKKWKGINKLPAITISMEKLNLPVIRSRAT-----RST 586
 Qy 1043 QNEPFURTPRNNTYPIKPATENWMSDPFLCTEWKASNSVPTVSHQLRPAKIVVAALGD 1102
 Db 587 QDGYLUISLO-----SLSISLDFT---KHFHSFPGOSTQLODGVS-FACKNN 631
 Qy 1103 SMTAVGARPNNSSDILPTSWRGLSWSTGGDGNLTHHTLPLNPKNNPY---LGFSTS 1158
 Db 632 TUTQKAKAAAQKSGGGS---SASCGSQGVFESNSTI--VLIVFGCIVCIVLIPAS 685
 Qy 1159 TWEGTAGINVVAEGARA-RDMPAQANDL-----VERMNKSPDINLRKDWKLVTLGEVN 1211
 Db 686 SWPYLHQRDPAARKCFARELPEMEEVKPVKKKDVSQMDNPKBERELBK----TOPBSIK 740
 Qy 1212 DICHYCENPBEAHATEVY-----QHIOQALDILSEELPRAFHV 1250
 Db 741 DVPKLLELDNSAAKTVDCRVEALQDMHKTCIPPLSEELJKKFGV 788

Search completed: January 6, 2004, 18:57:26
 Job time : 55 secs



Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Om protein - protein search, using sw model

Run on:

January 6, 2004, 18:33:20 ; Search time 370 Seconds

(without alignments)

790.635 Million cell updates/sec

Title: US-10-054-691-2

Perfect score:

7766

Sequence:

1 MGLRPGLFLLILLGQGT..... RCRGGGRREDPPMSLRTVAL 1458

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

733937 seqs, 200641211 residues

Total number of hits satisfying chosen parameters:

733937

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA,*

1: /cgn2_6/ptodata/1/pubbaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubbaa/PCTC_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubbaa/US06_PUBCOMB.pep:*

4: /cgn2_6/ptodata/1/pubbaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubbaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubbaa/PRCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubbaa/US07_PUBCOMB.pep:*

8: /cgn2_6/ptodata/1/pubbaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubbaa/US09A_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubbaa/US09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubbaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubbaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/1/pubbaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubbaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubbaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubbaa/US10C_PUBCOMB.pep:*

17: /cgn2_6/ptodata/1/pubbaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

19: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

20: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

21: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

22: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

23: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

24: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

25: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

26: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

27: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

28: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

29: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

30: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

31: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

32: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

33: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

34: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

35: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

36: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

37: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

38: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

39: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

40: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

41: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

42: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

43: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

44: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

45: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

46: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

47: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

48: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

49: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

50: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

51: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

52: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

53: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

54: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

55: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

56: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

57: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

58: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

59: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

60: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

61: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

62: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

63: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

64: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

65: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

66: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

67: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

68: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

69: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

70: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

71: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

72: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

73: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

74: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

75: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

76: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

77: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

78: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

79: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

80: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

81: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

82: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

83: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

84: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

85: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

86: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

87: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

88: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

89: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

90: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

91: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

92: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

93: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

94: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

95: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

96: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

97: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

98: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

99: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

100: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

101: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

102: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

103: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

104: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

105: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

106: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

107: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

108: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

109: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

110: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

111: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

112: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

113: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

114: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

115: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

116: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

117: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

118: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

119: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

120: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

121: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

122: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

123: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

124: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

125: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

126: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

127: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

128: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

129: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

130: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

131: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

132: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

133: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

134: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

135: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

136: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

137: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

138: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

139: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

140: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

141: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

142: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

143: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

144: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

145: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

146: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

147: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

148: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

149: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

150: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

151: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

152: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

153: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

154: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

155: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

156: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

157: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

158: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

159: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

160: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

161: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

162: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

163: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

164: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

165: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

166: /cgn2_6/ptodata/1/pubbaa/US60_PUBCOMB.pep:*

QY 181 PSAQQNGLAAGCVDLEMGVLDLQQEVRPAFNVLVDSEVAEVSRQHGTWLS PAPPCN 240
 Db 181 PSAQQNGLAAGCVDLEMGVLDLQQEVRPAFNVLVDSEVAEVSRQHGTWLS PAPPCN 240
 QY 241 CSBETTRAKVNWNSTOBANNILLASRSRYSBQESFTWVPPFYTTPSHLSEDERLOD 300
 Db 241 CSBETTRAKVNWNSTOBANNILLASRSRYSBQESFTWVPPFYTTPSHLSEDERLOD 300
 QY 301 STTLAMLNRMERPAGEKDELUSVKGRPMCPQSDESPLYLESYRNSNLTRQKOPKL 360
 Db 301 STTLAMLNRMERPAGEKDELUSVKGRPMCPQSDESPLYLESYRNSNLTRQKOPKL 360
 QY 361 EVREGABIRCPDKDPSDTVPPTSVRHKPADINVIGALGSITAGNGAGSTCPNVLDLTQ 420
 Db 361 EVREGABIRCPDKDPSDTVPPTSVRHKPADINVIGALGSITAGNGAGSTCPNVLDLTQ 420
 QY 421 YRGLSMSVGGDNIGNITVTANILRENPNSLKGFSVGTGKESPTNAFLNQAVAGGRAEDL 480
 Db 421 YRGLSMSVGGDNIGNITVTANILRENPNSLKGFSVGTGKESPTNAFLNQAVAGGRAEDL 480
 QY 481 PWQARRELVDLMNDTRHFQEDWKITLFIGNDLCFCNDLVHYSQNFTDINGRADI 540
 Db 481 PWQARRELVDLMNDTRHFQEDWKITLFIGNDLCFCNDLVHYSQNFTDINGRADI 540
 QY 541 LHEEVPRAFVNJTVLBTIUNBLBYQEKKVTCPRMILRSCLCPLVKEDDNSTELATIEF 600
 Db 541 LHEEVPRAFVNJTVLBTIUNBLBYQEKKVTCPRMILRSCLCPLVKEDDNSTELATIEF 600
 QY 601 NKKFOEKTHOLTESGRYDTREPTVUVQPFENNDMRTKSEGULPDNSFFAPCFHSSKS 660
 Db 601 NKKFOEKTHOLTESGRYDTREPTVUVQPFENNDMRTKSEGULPDNSFFAPCFHSSKS 660
 QY 661 HSRASAALWNMLPEPGOKTTHKFKENKINTICPNOVQFLRKYKSMQHGTWLPCCRDR 720
 Db 661 HSRASAALWNMLPEPGOKTTHKFKENKINTICPNOVQFLRKYKSMQHGTWLPCCRDR 720
 QY 721 ABSALHTPSVHLRPADIQWVALGDSITAGNGIGSKSPDDLVTTQYRGSLYSAGGDG 780
 Db 721 ABSALHTPSVHLRPADIQWVALGDSITAGNGIGSKSPDDLVTTQYRGSLYSAGGDG 780
 QY 781 LENVTTLNILEBFNRNLTGYAVGTGDAANDTNAFLNQAVPGAKADEIMSQVTLQKND 840
 Db 781 LENVTTLNILEBFNRNLTGYAVGTGDAANDTNAFLNQAVPGAKADEIMSQVTLQKND 840
 QY 841 DRVNFEHDWKVITVLLIGSDIDCYCTSDNL-SAANFVDLKNAVDLHREVPRVLNV 900
 Db 841 DRVNFEHDWKVITVLLIGSDIDCYCTSDNL-SAANFVDLKNAVDLHREVPRVLNV 900
 QY 901 DFLNPTIMQVFTGNDPKCPVQOASVLCNCVLTRENSQELARLEAFSRAVHSSMRLG 960
 Db 901 DFLNPTIMQVFTGNDPKCPVQOASVLCNCVLTRENSQELARLEAFSRAVHSSMRLG 960
 QY 961 SGRYTDODEFSVTLQPFONIQLPVLADGLPDTSFADDCIPNQKPHSOLARALWTML 1020
 Db 961 SKYDQDFSVTLQPFONIQLPVLADGLPDTSFADDCIPNQKPHSOLARALWTML 1020
 Qy 1021 ERIGSKETTDLRAEMPTCPONEPFLTRPRNSNTYPIKAENNGSDEFCTEWASN 1080
 Db 1021 ERIGSKETTDLRAEMPTCPONEPFLTRPRNSNTYPIKAENNGSDEFCTEWASN 1080
 Qy 1081 SVTTSVHQLRPAKIVVALGDSLTAVGARNNSDPLTSRGLSNSIGGONLEHTT 1140
 Db 1081 SVTTSVHQLRPAKIVVALGDSLTAVGARNNSDPLTSRGLSNSIGGONLEHTT 1140
 Qy 1141 LPNTLKKNPYLLGFSTSWEGTAGLNAAEGRARDNPAQADLVERMKNSPDINJEKD 1200
 Db 1141 LPNTLKKNPYLLGFSTSWEGTAGLNAAEGRARDNPAQADLVERMKNSPDINJEKD 1200
 Qy 1201 WKLUTLFIGVNDICHYCNPEAHLTAEYOHIOALDITSELPEPRAFNVMELASY 1260
 Db 1201 WKLUTLFIGVNDICHYCNPEAHLTAEYOHIOALDITSELPEPRAFNVMELASY 1260
 Qy 1261 QGQGKCAMLAQINCTCLRHSQSLSLEKLLKUWNLQHQISSLPSYWHQYOREDFAWV 1320

RESULT 2
 US-10-094-749-2029
 Sequence 2029, Application US/10094749
 Publication No. US20030219741A1
 GENERAL INFORMATION:
 APPLICANT: ISOGAI, TAGAO
 APPLICANT: SUGIVAMA, TOMOYASU
 APPLICANT: OTSUKI, TENSUJI
 APPLICANT: WAKAMATSU, AI
 APPLICANT: SATO, HIROYUKI
 APPLICANT: ISHII, SHIZUKO
 APPLICANT: YAMAMOTO, JUN-ICHI
 APPLICANT: ISONO, YUUKO
 APPLICANT: HIO, YURI
 APPLICANT: OTSUKA, KAZU
 APPLICANT: NAGAI, KEICHI
 APPLICANT: IRIE, RYOTARO
 APPLICANT: TAMECHIKI, ICHIRO
 APPLICANT: SEKI, NAOKO
 APPLICANT: YOSHIGAWA, TSUTOMU
 APPLICANT: NAGAHARA, KENJI
 APPLICANT: MASURO, YASUHIKO
 TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
 FILE REFERENCE: 084335/0160
 CURRENT APPLICATION NUMBER: US/10/094-749
 CURRENT FILING DATE: 2002-03-12
 PRIOR APPLICATION NUMBER: 60/350,435
 PRIOR FILING DATE: 2002-01-24
 PRIOR APPLICATION NUMBER: JP 2001-328381
 PRIOR FILING DATE: 2001-09-14
 NUMBER OF SEQ ID NOS: 3381
 SOFTWARE: PatentIn ver. 2.1
 SEQ ID NO 2029
 LENGTH: 270
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-094-749-2029
 Query Match 17.9%; Score 1392; DB 12; Length 270;
 Best Local Similarity 100.0%; Pred. No. 2.4e-119;
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 491 MNQDTRHFQEDWKITLFIGNDLCFCNDLVHYSQNFTDINGRADIILHAEVRAFV 550
 1 MNQDTRHFQEDWKITLFIGNDLCFCNDLVHYSQNFTDINGRADIILHAEVRAFV 60
 Db 551 NUVTULEVNBLBYQEKKVTCPRMILRSCLCPLVKEDDNSTELATIEFNNKFOEKHQ 610
 Qy 611 NLVTULEVNBLBYQEKKVTCPRMILRSCLCPLVKEDDNSTELATIEFNNKFOEKHQ 120
 Db 611 NLVTULEVNBLBYQEKKVTCPRMILRSCLCPLVKEDDNSTELATIEFNNKFOEKHQ 120
 Qy 611 LIESGRYDTREPTVUVQPFENNDMRTKSEGULPDNSFFAPCFHSSKSHSRAASAWN 670
 Db 121 LIESGRYDTREPTVUVQPFENNDMRTKSEGULPDNSFFAPCFHSSKSHSRAASAWN 180
 Qy 671 NMLEPVSGOKTTHKFKENKINTICPNOVQFLRTYKSMQHGTWLPCCRDRALSHTSV 730

RESULT 6
US-10-369-493-5416
Sequence 5416, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369, 493
CURRENT FILING DATE: 2003-02-28
PRIORITY APPLICATION NUMBER: US 60/360, 039
PRIORITY FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 6523
LENGTH: 348
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-369-493-6523

Query Match
Best Local Similarity 6.0%; Score 463.5; DB 12; Length 981;
Matches 118; Conservative 71; Mismatches 169; Indels 39; Gaps 14;

Query 358 DKLGVREG-AEICCPDKDPSDPTVTSVRLKPADINVIGALGAGSISLTAGNGAASTPGNLD 416
Db 43 DMSYPEPMGPFPFNCP--QPHRKSESSESVHOLHPSQIGVAAQGSVSVAQAKSS--SILD 98

Query 417 VLUOYRGLSWSGGDDENIGTIVTTLANILREFPNLSLKGFSVGTCGETSPTNAFLQNOAVAGR 476
Db 99 LFQDFPGPSFVTSPDVKAEOATFINERKFPIRK---GSDPFPVKYDFENMAIPCSF 1155

Query 477 AEDLPVQARRLVDLMKDNTTRINFODWKLTITLFTGGNDLCDFNDLVHSPQMTDNIGK 536
Db 156 SSELPEQALKLAKASTLKKKGRETSPTWDKVNIFIGHNDLNCINNETTFFGPERFGKSLSH 215

Query 537 ADLTHAEVPRAFNLVTVLETVNLRELYQEKRY--YCPMRMIRSLCPCVULKIDDNSTL 594
Db 216 ALTIITQTNVPKVFVN--IMPPINVKHSQAHKSKFC-BFSRKRTCSICIFELNEKEYQ- 270

Query 595 ATLIEFNKEQEKTQHOLE--SGYDPTRDFTVUQPFFENVPMPKTSSEGFLPNSFAFD 652
Db 271 ---NIKKQFDEQLENEVUVQFQKXGNSSTFPAVITAPANDLKSIP-LUKNQNPINGLLAD 325

Query 653 CFHSSSKSHRSASALWNMPLPEVGQKTRHKFENKINT---TCPNOVQPLRTYKNS 707
Db 326 CFHSPiANDIAAKQIKWGELEPDQKT---TINQLSVGFDRVCPVECPVLTIONS 381

Query 708 MQGIFTWLPCRDR---RAPSAHLPTSVALRADIOV 740
Db 382 ENCE---PSREMFLRVPSAFASSPGLSMPYMPWIPV 414

RESULT 7
US-10-369-493-6523
Sequence 6523, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369, 493
CURRENT FILING DATE: 2003-02-28
PRIORITY APPLICATION NUMBER: US 60/360, 039
PRIORITY FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 6523
LENGTH: 348
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-369-493-6523

Query Match
Best Local Similarity 5.2%; Score 406.5; DB 12; Length 348;
Matches 117; Conservative 66; Mismatches 133; Indels 47; Gaps 11;

Query 1060 IKPAIENNGSDFLCTEWKASNV-----PITSVHQLRPAKIVAAAGDSLTTAVGARN 1113
Db 15 VROMALE-----KYLNSIDPHRHYDPPDMVNMKHFHRVIGAMGDSL--IGSRAE 61

Query 1114 NSSDLPLPTSWRGWSLICGDGNLRHTHTPLNLIK---KFNPYLGFSSTWEG-TAGLN 1168
Db 62 NTIVGQRORYPGNMAFFTGMDFEVDRHLTVYNFRIAETKGKLFGGSTGIDYGENTGLN 121

Query 1159 AAEGARARDMPQAQDWJVERMONSPDTNLEKOMKLYTFIGND---LCHYCENPPEHLA 1225
Db 122 ATCGMKSDDILQAKENURSRKANTEENIENDKLUWSLWIGNDVGTGIGYRLEDPI--V 179

Query 1226 TEVQHQIQLAQDITISEBIPRATVNUVMEVMEALSYQOCGGKCAMLAQQNCTCLRHQS 1285
Db 180 DEVKSHTEKGGLYLUKENLPRPTIVSIVGMPADLQEAQ-----SILKOKRA 226

Query 1286 LEKOBELKCNWLNQHGLISSFSYTHQYQTR--EDFAVWQPR-FQNTLTPNERGDTDLT 1341
Db 227 RTVENOKLUD-DLSDGYNNSVTDQNNEHFNNSNDFTVWQPRATEY'TSYRDEHGKNT 285

Query 1312 PSEDCEPFSRDRGAEMAIAALNNMLPFGRKTSNFTHSPAKLCKPPSPSPYPLYRN 1401
Db 286 FYASDLEFLISKERCHAVLAKHYWNLNFPEVBGEKTRKADINGDTPKVIELNEKONLKITWGN 345

Query 1402 SRL 1404
Db 346 SKM 348

RESULT 8
US-10-369-493-6522
Sequence 6522, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369, 493
CURRENT FILING DATE: 2003-02-28
PRIORITY APPLICATION NUMBER: US 60/360, 039
PRIORITY FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 6522
LENGTH: 349
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-369-493-6522

Query Match
Best Local Similarity 4.5%; Score 347; DB 12; Length 349;

Best Local Similarity 31.9%; Pred. No. 1e-22; Mismatches 124; Indels 50; Gaps 11; Matches 106; Conservative 52; Mismatches 124; Indels 50; Gaps 11;

Qy 374 DSDPTVPSVHRIKPADINVIGALGDSUTAGNGA-----GSTPGNVLVDLVTQXQLS 425
Db 27 DPTVPSVHRIKPADINVIGALGDSUTAGNGA-----GSTPGNVLVDLVTQXQLS 75

Qy 426 WSGGGDENIGTVTLANIR---EFNPSLKGFSVGTKETSPNAFLNQAVAGGRAEDLP 481
Db 76 FFGIDESBDGHSIYNFRVIAETGKGFLGFGSTGVG-YGNTGLNVAVGGMKSDDIL 133

Qy 134 RQAKDLYNSRIKANEINKEWDKWLKVSLWIGTND---GNLVFGSENPIPVKYKAFEE 189

Qy 482 VARRLVDLIMKNDTRIHFQDKLITFIGNLICDFNDLWYSP---ONFTDIIKG 536

Db 190 GLYIYKLKNPLPRTIVSINGMPOLQBEY---YIATGTNRGTP---ESKOLDE 238

Qy 597 LIEFNKEQBKTHQIOLIESGRYDTREDFTVVQDF-FENNDMPKTSGLPDNIFRAPDCFH 655
Db 239 LD---SRNASEIQNQEGKFQDFE-FTVVQFGTEYDAFARNEFGNYSSALAYDVFH 294

Qy 656 ESSKSHSRASALAWNMLPVSOKTRIKFEN 687

Db 295 LGKLGQATVAKHYWQNLFPVSGERTKRAELGN 326

RESULT 9

US-10-369-493-5161

; Sequence 5161, Application US/10369493
; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slatter, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianteng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF FILE REFERENCE: 3B-10(2052)B

CURRENT APPLICATION NUMBER: US/10/369_493

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360, 039

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 5161

LENGTH: 425

TYPE: PRT

ORGANISM: Caenorhabditis elegans

US-10-369-493-5161

Query Match 1.8%; Score 141.5; DB 12; Length 799;

Best Local Similarity 19.0%; Pred. No. 0.042; Mismatches 270; Indels 255; Gaps 38; Matches 152; Conservative 125; Mismatches 270; Indels 255; Gaps 38;

Qy 405 NGAGSTPCBNVLDWLT---OYRGLWSVGGDENIGTVTLANIREFENPSLKGFSVGKE 461
Db 67 NNSSSKTENLNIFATSPREKDN-SVVG----NSTYHLASYL-----AFSLSHDLQ 112

Qy 462 TSPNAFLNQAVAGGRABEFLPVQARRLVDLKDTRIHFQDKLITFIGNLICDFND 521
Db 113 NSPKLISSEPM---KATGGLQIR-----DVFNS 140

Qy 522 LV---HYSPON---FTDNIGKALDIHAEVRAFVN-LVTVLB--INNLRELYQE-KVY 571

Db 141 LMVWVHHTSPQNYGMQALKKTSVMT-LKGADIDLWDOLEQVSVLMDFFKNISSVG 199

Qy 572 CPEMILSLSCPCVLFEDD--NSTELATLIEBFNKFOEKT-----HOL----- 611

Db 200 TGNLVNVNLVGLMEKAQSSHSWNVNHLQLSRLFPKDVDAVIDVYVLPHAVRLLQCV 259

Qy 612 ---IESGRYD-----TREDFTVVQDFPENVDMPK----- 638

Db 260 PGKNTITGGLDVKYTSFTLHGITSNITKEDFAVILDTTELVDSPKDDISEALACFP 319

Qy 639 -----TSEGFLDNPSPAPDCPH--FSKSHSRASALAWNMLPVGQ-----KTRRH 683
Db 320 VVWCWNHNTNSGRQNSKIDCPNCVHGLMSSSFKVVAISLDHFLSPCGEDSPCSNESSRM 379

Qy 684 KEPNK-----INTCPCNQVOPPELTYKNSMQGHWTWLPERDR 720

Db 380 ETRKWKVCIHELVDWMSILLESEVFVNNTSLVKAOKF-----WHKTFP 426

Qy 721 ABSALHPT---SYHALRDP-----IQVALGD---SLTAGNGICSSKPDLL----- 761

Db 427 VPNSINORDTSSELCGSGSIKQVAMQIIEKUKVNFTKVTSGENTILDKLSSLINKLIN 486

Qy 762 PDVTTQVRGLSYS-----AGGDSLLENVTLPNLREF---NRNLTGAVGTDAN 809

Db 487 EPTETSONIISNLERTVQLISEDMSLEKST--HNLSLFLMLQMNNTGSSL---E 539

Qy 810 DTNAFLNOAVPAGAKELMSOTLWOMOKDHVRNPHEDWKVITVIGGSDLCDYCTDS 869

Db 540 ALSFIERSETPYNFEEBLWPKFQOIMKDLTODFRIRH-----LL--SEMNGKIKI 588

Qy 870 NYUSAANFDHRLNALDVHREVPRLVNLVDFLNPTIMQVPLGNPDCKPQVQASVLCN 929

RESULT 11
 US-10-156-761-13312

Sequence 13312, Application US/10156761
 Publication No. US20030119018A1
 GENERAL INFORMATION:
 APPLICANT: OMURA, SATOSHI
 APPLICANT: IKEDA, HARUO
 APPLICANT: HORIKAWA, JUN
 APPLICANT: SHIBA, TADATOSHI
 APPLICANT: SAKAKI, YOSHIO
 APPLICANT: HATTORI, MASAHIRO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 243-262
 CURRENT APPLICATION NUMBER: US/10/156, 761
 CURRENT FILING DATE: 2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-204089
 PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO 13312
 LENGTH: 300
 TYPE: PRT
 ORGANISM: Streptomyces avermitilis

RESULT 12
 US-10-369-493-1156

Sequence 1156, Application US/10369493
 Publication No. US20030233675A1
 GENERAL INFORMATION:
 APPLICANT: CAO, Yongwei
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Slater, Steven C.
 APPLICANT: Goldman, Barry S.

RESULT 13
 US-10-359-012-4

Sequence 4, Application US/0359012
 Publication No. US20030232419A1
 GENERAL INFORMATION:
 APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
 APPLICANT: KOLODIN, Alex L.
 APPLICANT: TERMAN, Jon R.
 APPLICANT: MAO, Tianyi

APPLICANT: Chen, Xianfeng
 TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES
 FILE REFERENCE: 38-10(52052)B/10/369, 493
 CURRENT APPLICATION NUMBER: US/10/369, 493
 PRIORITY APPLICATION NUMBER: US 60/360, 039
 PRIORITY FILING DATE: 2002-02-21
 NUMBER OF SEQ ID NOS: 47374
 SEQ ID NO 1156
 LENGTH: 1708
 TYPE: PRT
 ORGANISM: Methanobacterium thermoautotrophicum
 US-10-369-493-1156

Query Match Best Local Similarity 1.7%; Score 133; DB 12; length 1708;
 Matches 127; Conservative 20.7%; Pred. No. 0.1; Mismatches 202; Indels 202; Gaps 29;

QY 593 ELATLIEBNKEQEBKT-HOLTESGRVDTREDFPTVVVQPFENMDPRTKSEGDPDNFFA 650
 DB 181 DLTSLIAVKQPKPGNTALHQWIDCALYAAAGKTNLNQFKLAKMVMYNNMGPQWSSWE 240
 QY 651 PDCFHFEKSRSRERASAMLNILPEVGQKTHHKFENKINTCPNOVOPFLR----- 702
 DB 241 P-----ATVEPASPISSEFYLYRDQFRFTKEDYFTYPLPAKPTWAVLSYVGSTGEV 292
 QY 703 TYKNSMOOGHTWLPCRIRAPSALHPTSHALRPAQVVAALGDSLTAQNGGSKPDLP 762
 DB 293 TYADAMO-----QITDALVSGLMVLPVIG--TWSNYY----- 323
 QY 763 DVTQTYRGQLSYAGGDGSLENVTPLNREF---NRNLTYGAVGTDANDTNAFLN- 816
 DB 324 -----ILNQTMQNIQIQTCPMPNOTYNTAIR-GIGNYDLSLIGV 364
 QY 817 QAVPGAKADELMSQVQTMKQMKDDHVNFDWKVITVULICGSDLCDYCTSNSLYSAN 876
 DB 365 TSVSSAN----VYEVQIL----DNGNLRLSLKISTV-----QPVNVIKA-- 400
 QY 877 FVDTHRMALDVHLREVRPVLMUDFLNPTIMQVFLGNGPKCPVQOASVUENCVLTRE 936
 DB 401 LVEKLPTASNVYQOE-----ANPEKYPV-KANVITDMLPTTG 437
 QY 937 NSOELARI-EAFSRAVRSRSMRBLVGSGRYT-----QEDFSVW----LQPFFFONIQ 982
 DB 438 STTSGASTTKFDRSNPFLRAMITSSTYRTIGQWIVSEEGPSWMSVWQACOPENQ-Q 496
 QY 983 LPTVLADGPDPTEFFAPDCIHPNQFHSQALARALWTNMELPLGSKTERLDLRAEMPTCTP 1042
 DB 497 IEPLAIGUE-----LQSDPTEP-GAQMDDITVTP 525
 QY 1043 QMEPP-----LRTPRNSN-----YTPPIKALENNSDPUCEWKAASNVPSTV 1086
 DB 526 RIEKLVSRFAFNIRLPMNSDKVVALVYNNP--PKQKNCASYL-----NPVSEI 575
 QY 1087 HOIRPADKVVALGDSLTTAVGARNNSSDIPTSW--RGWSIGGDNLETHTLPN- 1143
 DB 576 IEL---LKRMCAGTS---VGEIPQDADALVEMMIKGINVANWARGELEKLANSNA 627
 QY 1144 TLKFKNPVLGST 1157
 DB 628 ILMPYEDYLAWFNT 641

; APPLICANT: PASTERKAMP, Ronald J.
; APPLICANT: YU, Hung-Hsiang
; TITLE OF INVENTION: MOLECULES INTERACTING WITH CASL (MICAL) POLYNUCLEOTIDES, POLYPEPTIDE
; FILE REFERENCE: JHU1840-3
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/388,325
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 60/384,302
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: US 60/354,178
; PRIOR FILING DATE: 2002-02-04
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1633
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-359-012-4

Query Match 1.7%; Score 129.5; DB 12; Length 1633;
Best Local Similarity 19.3%; Pred. No. 0.2%; Mismatches 222; Conservative 118; MisMatches 424; Indels 385; Gaps 49;

QY 18 QGTPQIHTSPRSTLEQWLWPEITLKNSSPPCNPK-----LGVNMPSKSVHSLKPSD 69
Db 430 QGTPPLELAERESLY RLIQQT----PENINKNFEQYTLDPGTYKPNLNSHCVRPHQ 483

QY 70 IIEVVAAGNLIEPPDPGTDLEK-----ODWTERPOCYCMGMT- 108
Db 484 VKHLYTKELEYPLERLGSVRVSNLRSKEDIRPSKLTWCQQOTEYGYOHVNVTDLTT 543

QY 109 -----VLSDIRYFSTSV-----PRVPUCHGKRVPHD 136
Db 544 SRSGLACIATIIRFRPLINPFSNLNEADAVENNQAFDAVAREFGIPPVUTGKEMASQA 603

QY 137 GAEIDLWI-----OAQELVR-----NNKENIQLDFQPDWKUJINVFFSNASOCYL 179
Db 604 EBDKLSMWVYLSKFYELRGTPLRPWSWRKNGENDA-----L 644

QY 180 CSAQQNLAAAGCVDDELNGVLQDQEYVRAFYNLVD-----LSEVA 221
Db 645 AKSSISN-----NYLNLTPEKRTPRVDQGTCNDMNKKRKGFNTNLDEPS 690

QY 222 ESRQXHGTMWLPAPERNCSEITRLAKWVQWMSQYQBWANISLAS-----SRYSEQ 273
Db 691 NSSRSIG-----SNQCGSSKEGGNOKV-----ISMANQQLAKBESTRNPNSIMQK 739

QY 274 ESFPTVVFQPFYETT-PSLHSEDPRLOST-----TLAWHLMRMMERPAGE 318
Db 740 RVUSGIGKEPVULQSSGSPVHSCCPKPEATPSPPPLKQRFVVVTCVHL-RELKOVA 798

QY 319 KEPPLSVHGRMKPCPQESPYLFYSN-----SNYLTRLOKQDKLVE 364
Db 799 GBCCLS-----RPMRARAKSDSLQGGTENTPATPSTRPRAQALSGVLRQLOQEWKLQKR 854

QY 365 GAEIRCSDKDPTDVPSVHRLKPADINVIGLGSUTAGNAGSTGNVLDVLTQYRG 424
Db 855 ANQNLANREFHTPKNIKERKAH-----LASMEEHGDFEONKL---JSKGL 894

QY 425 SNS-----VGGDENITGVTTLANIREFNPSLKCFSVGIG---KETSPNAFLN 469
Db 895 STHHPPRPSRSPSPDPRASSPSPTVDS-ASPARKEKKPSOFPFHSHLRVHP--- 948

QY 470 QAVAGGRABDLPVQARRLVMDKNDTRHFQED-----WKLITLIGGNICDFNDL 522
Db 949 QLTVGKVSSGIGRAAEVLVNLVNDHRPKAQATSPDLESMRISPPNLLGGSDTCYFCKKR 1008

QY 523 VHVSPQNFDTNICKALDILHAEVPRAFMVLVTEIUNRELYQEKKVYCPMLRSCLP 582
Db 1009 VY-----VMERUSAEGHFRHRCFCSI-CATTIRLAYTCFDCDESKYCK-----P 1054

QY 583 CUVLFDDNINSTELLATEFFKKEQEK-THQLIESGRDVTREPVUVQOPPFENDMPKSE 641
Db 1055 HPIHCKTKNSKQCRKRAKIQDREEAATWQEQBAAPRDTPTSCAVAAI-----GTL 1107
QY 642 GLPDNSPFAFDPCDFHFSKSHSRAASAL-WNNMLE-----PVQGKTRHKFENKINTCPN 695
Db 1108 GPPVHRSLP-----VLUHPLGMILWDWNHSIPESTVHLKAGERISOKSAENGRRVULK 1160
QY 696 QYOPFLTYKNSMOGHGTWLFCDRDRAPSALHPTSVHARPAUDIOVALLGDS----- 747
Db 1161 PVRLL-----LP----RAAGEPLPTQORGAEQKMGTPAEAOQGERBNVPPRSP 1204
QY 748 -TAGNCIGSKDPLDPTVQRGLYSAGGD-GSLENVTTLPN-----ILRENRN 797
Db 1205 LRIANAIARRSLEPL-----LNSEGGKAWAQESKTLPTOACTRSFLRKTSN 1255
QY 798 LTGYAVGTDANDTNAFLNQAVPGAKAEDLMSQVQTLMKQMKDHRYVNHFEDWKVITVLU 857
Db 1256 KQG-----DQHSPGRNQSAFSPPDPLRTHSLPNIPS-----KVPAI- 1294
QY 858 GSSDLCDYCTDSNLYSAANFVUDHLRVALDVLRREVRLVLUVDEFNPTIMRQFEGN-- 915
Db 1295 -----RSPCCSKIEV-----PTLLEKVSQEN 1318
QY 916 PD-KCVQQAQSVLNCVLTURENSBLARJEAPSRAVSSRMREUFGGR-----YDTEQ 968
Db 1319 RDASKPKEKRISLFFS--LRKDKSPESFLOES--RQRKDIRDLCGSPKRVLPEDAQ 1373
QY 969 DFSVVIQPF 977
Db 1374 ALEKULQPF 1382

RESULT 14
US-09-808-602-84
; Sequence 84, Application US/09808602
; Patent No. US2002015115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: Macdougall, John
; TITLE OF INVENTION: No. US2002015115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 1566-697.CTP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 2765
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-808-602-84

Query Match 1.7%; Score 128.5; DB 10; Length 2765;
Best Local Similarity 17.9%; Pred. No. 0.66; Mismatches 277; Conservative 194; MisMatches 51; Indels 565; Gaps 71;

QY 256 SYQABAMSLASSRYSEQESTVVFOPPYETPSLHSEDR-----LQDSTLWHAH 307
Db 1112 AVTFINDKTDAYGTVQGLSDBAVSVGEYETCPSLJLWEKTLQGELFEDPSNUGGS 1171
QY 308 LMNRMMEPAGEKDEPLISWKG-----RPNKCP----- 335
Db 1172 L-----DGKHTLNVKSGILLKGGENOPLTOOPAIIITSIMGNRRSISCPSCNGL 1222

QY 336 QESPYLFS-----YRNSNYLTROKQD----KLEVERECAETRCPDF--- 373
 Db 1223 AEGNKLLAPVALAVGIDSSLFVGDFNYKRRFFSRNVTISILERNKEFKHSNSPGHKYI 1282
 QY 374 -DP-----SDTPVPTSVHRLKPADINVIGALGDSLTAAGNAGSTGCVNLVLTQRL 424
 Db 1283 AVDPVTGSLYVSDTNSRRYRVK-----SL---SGAKDLAGNSEV----- 1320
 QY 425 SWSVGGDENIGNVTITLNLREFNPNSLKGKPSVGTGKETSPNAFLNQAVA--GGRAEDLPV 482
 Db 1321 -----ACTGEOCLP---FDEARCDDGSKAVATL 1346
 QY 483 QARRLVLDMKND----TRHRQEDWKKITFIGGNDL---CDFCIDLVHVSPO 528
 Db 1347 M\$TRGIJAVTDKNGLMYFUDATMIRKUDQNGIISIULGSULATVRPLSCDSMVAQTRL 1406
 QY 529 NFTDTNGKALDLHAEVRAFRALVLTVE--VNRLRELYQ---EKKVYC----- 572
 Db 1407 WPTDLAVNPMD-----NSLYLVENNWLIRITENHQVSIATGRPMHCQVPGIDY 1455
 QY 573 -PRMLILSCLCPVLCFKFDNSTELATLIFPNKQKOBKTHQOLIESGRY-----DTRE 621
 Db 1456 LSKLAIRHAELESASAIAHSHTSYLITTEDEKKINRLQVTTINGECILLAGASDCCKN 1515
 QY 622 DFTVVVQPFEN-----VDMPKTSSEGFLPDNSIFAPDCPFHFSKSHB----- 663
 Db 1516 DVNCICYSGDDAVATDAILNPSLSSLAVALDGTIVIADGNIRRAVSKGKPVLAFAKQY 1575
 QY 664 AS-----ALWNNNLEPVEQTKT---TRHFKENKNITCPNQVOPFLRTKNSMOGHGTLW 716
 1576 AASGEQELLYVFNADGTHQYTSVLSLTGVEXLYNFTYSAADNVTELIDNINGSL----- 1627
 QY 717 CRIRAPSALHPTSHALRPADIQVAALGSDTAGNGGSKPDLPDTQ-----Y 768
 Db 1628 -KIRRDSEGMPR--HLMPP-DNOIT-----TATVGTNGGLKA-----VSTOULEGLMTY 1673
 769 RGLS-----YSGGDSLLENVT-----TLPNLLREFRNLTGAVGTDAN 809
 Db 1674 DSGWGLLATKSDETGWTTFYDVPHEGRITNVRPTGVUTSLAREMERSITIVIENSNRD 1733
 QY 810 DTNAFLNQAVPGAKAEDLMSQVO-----TLMQKMKDDHVNFB----- 848
 1734 DTVTIVTINLSSEASYTVVQDQVRNSYOLGSNGTLRVMYANGKQVFHSPEPHVLAGLPT 1793
 QY 849 -----DWKV-----ITVL-----IGGSDLCDXTSDMSIYSAANFV 878
 Db 1794 IGRCNISLPMENGANSIWRKREBOKIGKVTIGKLRVHGRNLSIDYDRITERKYD 1853
 QY 879 DHLENALDVLHREYPR-----VLVNLVDFLNPTIMQVFLGNPDKCPVQQAQLC 928
 Db 1854 DHKRFTRLTYDVQGRPEFLWLOSSGLAAVNVSFFNGRLLAGIORGANSERTIDKQG--- 1910
 QY 929 NCVLTRENSQLEARLEAFSRAY-RSSMRELWASGR-----YTOEDPSVWVQ----- 976
 1911 -----RIVSRMFADGKWSYSLDKSNVLLQSOQVIFEFDSDRHLHAVTMPSVARHS 1964
 QY 977 -----FFQNTOLP-----VLADGP-----TSFAPPOTHPKQFH----SOLA 1012
 Db 1965 MSTHSTSIGTRNTYNNPESNASTIFDXSDGGRILKTSLGTG---RQFYXKGKUSKL 2020
 QY 1055 NYTY-----PIKPAI-----ENWGSDFLCTEWKASNSVPSVH 1087
 1013 RALM-----TNMLEPLGSKTEFLDRAEMPTCPCIONEPFLTRPNS----- 1054
 Db 2021 ELYVYSTAVTFGYDETTGVLKMNQLQSGFCSTIRYRKVGPPVDQ1QYRFSSGMINARF 2080
 QY 308 LMNRMPFAGEDEPLSKHG-----RPMKCP----- 335
 1112 ATYFIWDTKDAYQVRVGLSDAWSVGFEYETCPSLILWEKRTALLQGFELDPSNLGWS 1171
 QY 326 QESPYLFS-----YRNSNYLTROKQD----KLEVERECAETRCPDF--- 373
 Db 1223 AEGNKLLAPVALAVGIDSSLFVGDFNYKRRFFSRNVTISILERNKEFKHSNSPGHKYI 1282
 QY 374 -DP-----SDTPVPTSVHRLKPADINVIGALGDSLTAAGNAGSTGCVNLVLTQRL 424
 QY 1098 QL-----RADIKV-----VALGSDUTAV-----GARNNSDLPTSWRG 1124
 Db 2140 TLSKHFDTGTRIKEVQYQVMNFRSLMYMWMTWQYDNGRVKRELKLGPYANTK----- 2191
 QY 1125 LSWSIGGDNLLETHTLFNUILKKFPNPLLGFSSTWREGTAGIN---VAECAARDMPA 1180
 Db 2192 YTDVQDGDQLO-----SVATNDRP-----TWSYDQDNLHLLNFGNSARLMP 2237
 QY 1181 QWDLVERMKNSPDINIEKD----- 1210
 Db 2238 R-YDLRORITRIGDVQKIDDGYLCOGRGSLIFEYNSKGULLTRAYNQASGSNSVQRYDG 2296
 QY 1211 NDLCHYCENPEAHLATY--YHOIQAIDLSEELRAFAVNU--EVNEASLY----- 1260
 Db 2257 SRRASYKINLGLHQFYSDLHR-----PTRITHVYNHSENSETISLYDQGH 2344
 QY 1261 ---QGOCGKCAMLAQNCITLRRHOSSLEKQELKVNWLQHGSFSYWHQYQRRED 1316
 Db 2345 LRAMESSSGEETYVASINTGTL--AVFSINGLMIKOLQYTAQE-----YYPNSPD 2395
 QY 1317 FAVVQPFQONTLPLNERGDTLTFSSDCDFHSORGHAEMAIALWNMLPVGKITS 1376
 Db 2396 FQMV-GFHGGIYDPLK-----LVHTORDYDVL-----GR-WTS 2430
 QY 1377 NNETHHSRALKCPSPSPSYLTYLRSRLLPDOE--EAEPEVLYAW 1420
 Db 2431 PDYMMWRNGKEPAPEN-LYMFKNKNPLSNELDKNVYVTDVSKWL 2475

RESULT 15
 US-09-800-198-72
 ; Sequence 72, Application US/09800198
 ; Publication No. US20030087816A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Vernet, Cornie AM
 ; APPLICANT: Fernandes, Elma
 ; APPLICANT: Shirkets, Richard A
 ; APPLICANT: Hermann, John L
 ; APPLICANT: Majumder, Rumud
 ; APPLICANT: Mishra, Vishna
 ; APPLICANT: Meza, Peter S
 ; APPLICANT: Rastelli, Luca
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: 15966-697
 ; CURRENT APPLICATION NUMBER: US/09/800,198
 ; CURRENT FILING DATE: 2001-03-05
 ; PRIORITY APPLICATION NUMBER: 60/186,596
 ; PRIORITY FILING DATE: 2000-03-03
 ; NUMBER OF SEQ ID NOS: 98
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 72
 ; LENGTH: 2765
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 ; US-09-800-198-72

Query Match 1.7%; Score 128.5; DB 11; Length 2765;
 Best Local Similarity 17.9%; Pred. No. 0.66; Mismatches 511; Indels 565; Gaps 71;
 Matches 277; Conservative 194; MisMatches 511; Indels 565; Gaps 71;

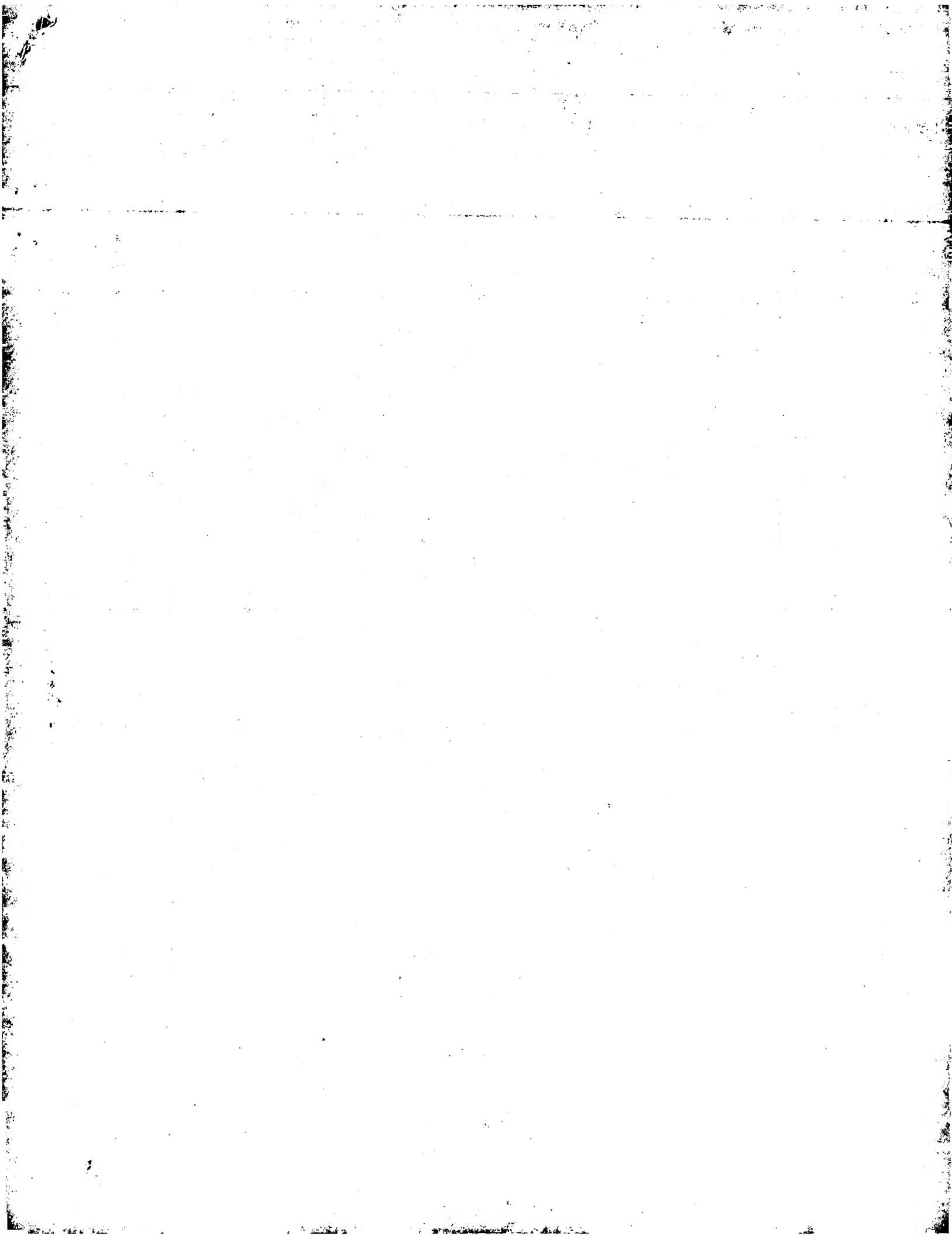
QY 256 SYQBAMNNSILLASSRYSRQESFTWVFOFFPYETPSLHSEDPR-----LQDSTTLAWH 307
 Db 1112 ATYFIWDTKDAYQVRVGLSDAWSVGFEYETCPSLILWEKRTALLQGFELDPSNLGWS 1171
 QY 308 LMNRMPFAGEDEPLSKHG-----RPMKCP----- 335
 1112 ATYFIWDTKDAYQVRVGLSDAWSVGFEYETCPSLILWEKRTALLQGFELDPSNLGWS 1171
 QY 336 QESPYLFS-----YRNSNYLTROKQD----KLEVERECAETRCPDF--- 373
 Db 1223 AEGNKLLAPVALAVGIDSSLFVGDFNYKRRFFSRNVTISILERNKEFKHSNSPGHKYI 1282
 QY 374 -DP-----SDTPVPTSVHRLKPADINVIGALGDSLTAAGNAGSTGCVNLVLTQRL 424
 Db 1283 AVDPVTGSLYVSDTNSRRYRVK-----SL---SGAKDLAGNSEV----- 1320

QY 425 SWSVGGDENIGNVTITLNLREFNPNSLKGKPSVGTGKETSPNAFLNQAVA--GGRAEDLPV 482

Db 1321 ----- ACTGEOCLP----FDEARCDDGDKAVATL 1346
 Qy 483 QARRLUDLUMKND-----TRHFQEDMKITLFIGNDL----- CDFCDLWHSPO 528
 Db 1347 MPPRGIAVDKNGLMYFVYDATMTRKVUDONGITLTGSDNLTAVRPLSCSDSNMDAQRLE 1406
 Qy 529 NFTDNIGKALDILHAEVPRAFNLVTLIE--IVNLRELYQ----EKKVYC----- 572
 Db 1407 WPTFLAVNPMD-----NSLVLENVNVLRITEHNPQSIAGRPMICQVPGIDY 1455
 Qy 573 -PMLWLNSLCPVCYLKDNNSTELATLBFNKHQEKHOLISGRY-----DTRE 621
 Db 1456 LSKLAIHSALESASAIAISHTSVLYITBDEKKINRLRQVTNGEICLLAGASDCCKN 1515
 Qy 622 DTTWVWOPPFEN-----VDMPKTSEGULPDNEFAPCOPHESSKSHR----- 663
 Db 1516 DYNICCYSGDDAVATDAILNSLSLAVADPDTIYIADLGNITIARVEKNPKVNAFOYE 1575
 Qy 664 AAS---ALWNNMLEPVQOKT---TRHKFENKINICPENOVOPLRTYKNSQGHGTWLP 716
 Db 1576 AASPGEQELYVNPADGHQYTSLVTGELYLYNTSYADNDVTEBLIDINGNSI----- 1627
 Qy 717 CRDRAPSALHPSTSVAHLRPADQWAAALGDSUTAGNGGSKDDLPPTO-----Y 768
 Db 1628 -KIRRDSSGMPPR--HLLMP-DNOII-----TIVUGTNGLKA-----VSTOMBLGLMTY 1673
 Qy 769 RGLS-----YSAGGDGSELENVT---TLPNLRPBNRNLTGAVGTDAN 809
 Db 1674 DENTGULLAKTSBETGWTTFYDHDREGLTNUVPTGVUTSRMEMSITUDENSIRDN 1733
 Qy 810 DTNAFLNQAVPGAKADELMSQO-----TLMQKWDODIRVNFHE----- 848
 Db 1734 DVIVITINLSSVASYTVQDQVRNSYQCLSNGSLTRVMYANGKGSFISEPHVLAGLPT 1793
 Qy 849 -----DWKV-----ITVI-----IGSDLCDPCTDSNLYSANFV 878
 Db 1794 IGRCNISLPMENGINSIEWLRKEQIKGKVTFGRKLVRHGMRNLSIDYDRNIRTEKYD 1853
 Qy 879 DHURNALDVLUHKEVPR-----VUVNLDPLNPTIMRQVFLGPDKCPVQQASVLC 928
 Db 1854 DIRKFTTRIYVQGRPEFLWUPLLSSGLAUNVNYFFNGERLAGJQORGANSERTDIDKOG-- 1910
 Qy 929 NCVLTRENSQBLARLEAFSRAY-RSSRRELVGSR-----YDQEDFSVWQ----- 976
 Db 1911 -----RIVSRPFDAGKWSYSLDKSNSVLLQSQROQYIFEDSSDRHLHAVTMPSVARHS 1964
 Qy 977 -----FFONIQOLP-----VLADGLPD-----TSPFAPCIPHIQKFH-----SOLA 1012
 Db 1965 MSTHTSIGYIRUNYNNPPESNASVIFDXSDDGIRLKTSFLGTC---RQVFVYKGKUSKLS 2020
 Qy 1013 RALW-----"TMMLEPLGSKBTBLDRAEMPICTCPQNEPFLTPRN----- 1054
 Db 2021 EIVYSTDSTAVFGYDGETGVLKMNLOGGFSCHIRYKVGPRVYDQJYRFSEGMINARF 2080
 Qy 1055 NYTY-----PIKPAT-----ENWGDDFLCTEWKASNSVPTSVH 1087
 Db 2081 DTYVHDNSFRASIKPVSETPLPVLYRDIEISGKVBHFK-FGVYIYDINQITAVM 2139
 Qy 1088 QL-----RADIKV-----VAALGHSITTAV-----GARPNNSDLPTSWRG 1124
 Db 2140 TISKHFDTGTRGRIKEVOYENFRSLNMWMVQYLSMGRVTKREKLGPYANTIC----- 2191
 Qy 1125 LWSIGGDNLEHTTIPNLKKENPNVPLGFSTTWEGTAGIN---VAEESARADMPA 1180
 Db 2192 YRYDYPDGDQLO-----SVANDRP-----TWSYDINGNLHNPNSARAMPL 2237
 Qy 1181 QAWDLVERMKNSPDINLEKD-----WILVUTLFIGV 1210
 Db 2238 RYDLRORITRGLDVQYKIDDGQYLQRGSDFEYNSKGLLTRAYNGASGSQVYRDGV 2296
 Qy 1211 NDICHYCNPEAHATEY---VORHQADLISBELPRAFYNTV--BYMELASSY----- 1260

Db 2297 SRRASVYKTNLGHQLQFYSIDLHH-----PTRLTHVNHNSBEITSLYVYDQGH 2344
 Qy 1261 ---QSGGKCAMLAQONCTCLRHSOSSUBKOELKVNWLQHGSSFSYWHQYTORRED 1316
 Db 2345 LFAMESSSGBEYVVASDNTGIPB-AVSFINGLMIQOLQYTAGET-----YDSNPD 2395
 Qy 1317 FAVVQPFQNLITPLNERGOTDLTFSOCFHFSORGHAEMAIALWNMMLBPVGKTS 1376
 Db 2396 PQWVI-GFHGGLYDPLK-----LVHFTQDRYDVA-----GR-WTS 2430
 Qy 1377 NNFTHSRAKLKCPSPESPSYLTLRSRNLPPDAB---EAEVLYWAV 1420
 Db 2431 PDYTMWRNGKEPAPFN-LYMFKNKNPLSNEBLDKNVNTDVKSMLV 2475

Search completed: January 6, 2004, 19:05:43
 Job time : 379 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

Om protein - protein search, using SW model

Run on: January 6, 2004, 18:59:25 ; Search time 31 Seconds
(without alignments)

Title: US-10-054-691-2
Perfect score: 7766

Sequence: 1 MGLRPGIFLLELLLUGQT..... RCRGGGRREDPPMSLRTVAL 1458

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep: *
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep: *
5: /cgn2_6/ptodata/2/iaa/PCTR_COMB.pep: *
6: /cgn2_6/ptodata/2/iaa/backfile1_pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

* Query

Result No. Score Match length DB ID Description

| | | | | | | | |
|----|-------|-----|-------|-----------------|----------------------|--|--|
| 1 | 138 | 1.8 | 10182 | 4 | US-09-134-001C-3159 | Query Match Best Local Similarity 18.4%; Pred. No. 0_1; Matches 265; Conservative 187; Misnmatch 546; Indels 440; Gaps 61; Sequence 3159, Ap | US-09-134-001C-3159 ; Sequence 3159, Application US/09134001C ; GENERAL INFORMATION: ; APPLICANT: Lynn Doucette-Stamm et al ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS ; FILE REFERENCE: GRIC-007 ; CURRENT APPLICATION NUMBER: US/09/134, 001C ; CURRENT FILING DATE: 1998-08-13 ; PRIORITY APPLICATION NUMBER: US 60/064, 964 ; PRIORITY FILING DATE: 1997-11-08 ; PRIORITY APPLICATION NUMBER: US 60/055, 779 ; PRIORITY FILING DATE: 1997-08-14 ; NUMBER OF SEQ ID NOS: 5674 ; SEQ ID NO 3159 ; LENGTH: 10182 ; TYPE: PRT ; ORGANISM: Staphylococcus epidermidis |
| 2 | 136 | 1.8 | 1455 | 2 | US-08-726-0121-2 | Sequence 2, Appli | 4555 TIRKNAKDNLNSMTALSESIADYENQKQENYLDASNNKRQDYDNAAKGILNQTOQP 4614 |
| 3 | 129.5 | 1.7 | 1218 | 4 | US-09-589-567-2 | Sequence 2, Appli | 412 GNVLDVHQ---YRGLSWVGGDENGITVTLA---NLRFENSLKGFSVGKETS 463 |
| 4 | 127.5 | 1.6 | 1016 | 3 | US-09-180-439-8 | Sequence 2, Appli | 4615 TMSADVDTQAKDVKTTKAJDGNOLLEVAKQOALHNLTNDLAQRLTDTNH-S 4673 |
| 5 | 123.5 | 1.6 | 2636 | 4 | US-09-252-991A-25753 | Sequence 2, Appli | |
| 6 | 122 | 1.6 | 855 | 4 | US-09-107-532A-5646 | Sequence 2, Appli | |
| 7 | 121.5 | 1.6 | 1112 | 3 | US-09-353-585-2 | Sequence 2, Appli | |
| 8 | 120.5 | 1.6 | 1 | US-07-741-940-7 | Sequence 2, Appli | | |
| 9 | 120.5 | 1.6 | 2842 | 1 | US-08-289-548A-7 | Sequence 2, Appli | |
| 10 | 120.5 | 1.6 | 2842 | 1 | US-08-452-654-7 | Sequence 2, Appli | |
| 11 | 120.5 | 1.6 | 2842 | 4 | US-08-442-731-7 | Sequence 2, Appli | |
| 12 | 120.5 | 1.6 | 2843 | 1 | US-07-741-940-2 | Sequence 2, Appli | |
| 13 | 120.5 | 1.6 | 2843 | 1 | US-08-289-548A-2 | Sequence 2, Appli | |
| 14 | 120.5 | 1.6 | 2843 | 1 | US-08-452-654-2 | Sequence 2, Appli | |
| 15 | 120.5 | 1.6 | 2843 | 1 | US-08-452-654B-2 | Sequence 2, Appli | |
| 16 | 120.5 | 1.6 | 2843 | 1 | US-08-452-655B-7 | Sequence 2, Appli | |
| 17 | 120.5 | 1.6 | 2843 | 2 | US-08-370-238A-2 | Sequence 2, Appli | |
| 18 | 120.5 | 1.6 | 2843 | 3 | US-08-450-582-2 | Sequence 2, Appli | |
| 19 | 120.5 | 1.6 | 2843 | 3 | US-08-450-582-7 | Sequence 2, Appli | |
| 20 | 120.5 | 1.6 | 2843 | 4 | US-08-442-731-2 | Sequence 2, Appli | |
| 21 | 120.5 | 1.6 | 2973 | 2 | US-08-821-355A-7 | Sequence 2, Appli | |
| 22 | 120.5 | 1.6 | 2973 | 2 | US-09-003-687A-7 | Sequence 2, Appli | |
| 23 | 120.5 | 1.6 | 2973 | 3 | US-09-136-605-7 | Sequence 2, Appli | |
| 24 | 120 | 1.5 | 15281 | 2 | US-08-471-119A-2 | Sequence 2, Appli | |
| 25 | 119.5 | 1.5 | 1112 | 3 | US-09-353-585-3 | Sequence 2, Appli | |
| 26 | 118 | 1.5 | 1007 | 3 | US-08-961-082-216 | Sequence 2, Appli | |
| 27 | 118 | 1.5 | 1007 | 4 | US-09-536-784-216 | Sequence 2, Appli | |

QY
 464 PNA-FLNOAVAGGRADLPVQARVLUDLMKIDTRHFQEDMKITLEFIGGNDLDFCND 521
 QY
 4674 PNINSNVSNQADE-KANTVNTAMTQLQKOTIANYDDELLH---DENVN--ADKDKDAYN 472
 QY
 522 LVHYSPOPNFTDNIGKALDLHAEVPRAFVNLYTVLETVNRLRELYOEKKVYCPRMILRSLC 581
 QY
 4726 AVNNIKQLINQSDAMQAOLPAREINKVTORINTTKINDLNGNDKLAEK----- 477
 QY
 582 PCVLFKFDIESTELATLIEFNMKFOEKTHOLIESGRVDTREDPTWVQPPFE-NVDMPKTS 6400
 QY
 4774 -----RDANTTIDGTYLNAQRNKAKENY--GAKSTKNTSLODQYNQNMIAQMQLR 482
 Db
 641 EGLPD-----NSFFAPDCFFSSKSHSRAASALWNMMEP-----VGQKT--- 680
 QY
 4826 NSVNUVVNVKANSNTINEDNGPKEAQNQAVTH--AOTLINQSNPMSRDVWNQKQAV 488
 QY
 681 -TRH-----KFE-----NKINIT-----CPNQVOPFLRTYK 706
 Db
 4883 NTAAONLHGQOKLEAQOSSANTEIGNLPLNTQKAKEKELVNSKQTRTEVQBLNQAKS 494
 QY
 707 SMQSGHTWPCDRPALSALPTSHALRPADIQVVALGDSLTAGNGISK_PDLDPVT 765
 Db
 4943 LDSSNGTLKSLSVAKOPT-VOKTSVYNTEDOPQ---SAYNDSITMGOTLINKYADPVLDKT 499
 QY
 766 TOYGLSYSAGGGDSLENVTPLNLIREFRNRLTGVAVGTDANDTWAFLN---QAVPG 821
 QY
 5000 LV-----DNAISNISTKENALHG-EQKLTTAKTEAINALNTLADLNTPQEAKT 504
 QY
 822 A-----KAEDIMSQVTLMKQKDDHVNFHEDWKVITVLUFGSDLCDY 866
 Db
 5049 AINTTAHTRDVTAECOSKANGINSAMTTRONISDNESVT----- 508
 QY
 867 TDSNUYSAM---NFUDHRLNALDULHREPVRLVYDLEFINPTMRQVFLGNPDKCPV 921
 QY
 5088 NESNTINAPEKEQHAFTEALNNAKEIVNEODATLANSIN----- 512
 QY
 922 QOASVLCNCVLT-----RENS-OELARLEAFSRAYASSMRELVGSGRYDQT--- 967
 QY
 5128 QKAOQALITTKNALDGEQEEQLRAKENADQEINTLNQLTDAORNSEKGJLNSQTRTEVASQ 518
 QY
 968 ----EDFSVYLQPFEPFQNTQIPVLAGD---LDTSFAPDCHPNPKHSQLBALWTML 102
 Db
 5188 LAKARLNKYM-----QNLHLINGKNOMINSSKFNEDANOOQAYNSNALA----- 523
 QY
 1021 EPLGSKTETDLRAEMPICTPQNPEPLRTPRNNTYPIKALENNGSDFLCTEWKASN 108
 QY
 5236 EALKNKSQNPED-----KTYIE-----QAIN 525
 QY
 1081 SVPTSVHQLRPAKIVAAALGDSLTTAVGARPBNNSDPLTSWRGLSNSIGGDQNLEHTT 114
 Db
 5258 NINSAINN-GEAKLTKADEV-----ASINNLSGL----- 528
 QY
 1141 LPNTLKKENPYLLGSTSTWEGTAGLNAARGARDMPAQAW----- 529
 Db
 5290 -----TNBOKTKENQAVNGA-----DIV 118
 QY
 1187 ER-----MKNSPDINLKEKDMLVLTFIGVNDLHYCENPEAHLT---BY 122
 Db
 5332 NNQWIAHSTSNSYFNNDSTQNTYDWDINGSYTT---GOH-----NPELKNTIDT 538
 QY
 1229 VOHTQALDITL-BELPR-----AFVNVUEVMELASLYOCOGKCMALAO 127
 QY
 5382 ISRINTAKONDLHGVEKIQDQKGTAQNEBIGOLGYLNDPQKSGEESLVSNTSEVEHIL 544
 QY
 1275 NCTCLRHQSLSLEKPKKVNWLQHGISSSYWHO-YTQREDFAVWQPFQNTLPL 133
 QY
 5442 EAKSINNAMKOLDRDVKVAKTNUVKQSSDYINDSTEHQGRYDOR-----LOEAENI 549
 QY
 1333 NERGDTDL-TFFSDFCFHFSDRGHEMAIHW---NNMLPFPGRKUTSNNTHSRA 134
 QY
 5492 NEGNPNTLNKSEIBOKQOOLTAQNAOLQGSHLBEAKNATGINKLITALNIAQRQKA 554
 Db

RESULT 2
US-08-726-012B-2
Sequence 2, Application US/0872601B
; Patent No. 5932190
; GENERAL INFORMATION:
APPLICANT: Hans Joenie, et al.
TITLE OF INVENTION: CDNA FOR FANCONI ANEMIA C
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klarquist Sparkman Campbell Leigh
STREET: One World Trade Center, Suite 1600,
CITY: Portland
STATE: OR
COUNTRY: USA
ZIP: 97204-2988
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk, 3.5-inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: WordPerfect 5.1+, ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726, 012B
FILING DATE: 10/04/96
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Richard J. Polley
REGISTRATION NUMBER: 28,107
REFERENCE DOCKET NUMBER: 3812-45520/RJP/DJEE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1455
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: Linear

| Query | Match | Best Local Similarity | Score | Length |
|-------|--|-----------------------|-------------------------|--------------|
| Qy | 19.3% | Score 136; DB 2; | 1.8%; Score 136; DB 2; | Length 1455; |
| Db | 143 IQAQEJVRNKENLQDQFDWKIJNFFNSNASOCYLCPSAQONGLAGGVDBLMG--- | 198 | 19.3%; Score 136; DB 2; | Length 1455; |
| Qy | 366 LSAEELVGHQEVLEQ-EHWFQWRLSFTSALVCF--PEAQO--LLDWARLMAQAFE | 420 | 19.3%; Score 136; DB 2; | Length 1455; |
| Db | 199 -----VUDYLOQEVPRAFVNVLDSVAEVS-ROVHGWTISPAPECNCBET | 245 | 19.3%; Score 136; DB 2; | Length 1455; |
| Qy | 421 SCQLDSSMVTAPFLVVRQALEGGPSAFLSYADWPKASFGSTRGHG-----CSKKKA | 469 | 19.3%; Score 136; DB 2; | Length 1455; |
| Db | 246 -----TRLAKUVMQHSIQEAWNSSLASSNYSEQESPTVVFQPPYET--TPSLHSEDPRL | 298 | 19.3%; Score 136; DB 2; | Length 1455; |
| Db | 470 LVLFPLFPLSELVP-----FESPYLQVHLHPBLPVSKYRSLTDYSLAKTRRL | 518 | 19.3%; Score 136; DB 2; | Length 1455; |
| Qy | 299 QDSTIAWHWN-----RMMPAGEKDEPS-----VKG-----GRPMKCPSQEPEL | 341 | 19.3%; Score 136; DB 2; | Length 1455; |
| Db | 519 AD---LKVSLEENMGXIEDLSSAGITEPHQSALQDVERKAIMVBEHTGIPVTTMEASIFR | 575 | 19.3%; Score 136; DB 2; | Length 1455; |
| Qy | 342 FSYRNNSYLTRIQLQPKDQKLEVREGAEIRCPDKDPSDPTVTSVHLR-----KPADI | 391 | 19.3%; Score 136; DB 2; | Length 1455; |
| Db | 576 RPYVVSHFLPALITPRVPLKPVDPDSRAFIESLKRADKIPPSLYSTYCOACSAABEKPED | 634 | 19.3%; Score 136; DB 2; | Length 1455; |
| Qy | 392 NVIGALGDSLTAGNSAGSTPGNVIDVLTVORGLGSWSVGDDENIGTVTTLANIREFBNFNSL | 451 | 19.3%; Score 136; DB 2; | Length 1455; |
| Db | 635 -----AALGVRRAEPNS-----AEEPLQQLTAALGEI----- | 661 | 19.3%; Score 136; DB 2; | Length 1455; |
| Qy | 452 KGFSGVGTGKETSPN-----APLNQ---AVAGGAEDLPYQARRVVDLMDNTRHF | 499 | 19.3%; Score 136; DB 2; | Length 1455; |
| Db | 662 -----ASWMDPSODRVISAQVAVISETSELRULAVIGHNEEDDSVSBISK----IOLSINTPRIEP | 713 | 19.3%; Score 136; DB 2; | Length 1455; |

; ORGANISM: Zea Mays
US-09-589-567-2

WY
Db 714 REHAVDIL-----LTSFPCQNLMAAS-----SVAP----- 738
559

OY
560 NLRBELYQEKKVYCCPRMILRSCLC---PCVLFDDNSTELATTIEFNKKFOEKTHOL--- 611

Db 739 -----PERQCPWAALFVRTMCGRLPAVL-----TRLCQQLRHQGPSLSSAPVHTLGLAA 786

| | | |
|----|---|----------------|
| Dy | 612 --IBSGRYPRDEFTVWQPFENNMDPKTSBGLP----- | BNSFAPDCF. 654 |
| Db | 787 LAVFGL--EARSALPYV----DYGEPAPGAGLPVPALEDDSLITCRTRDSLFE-- | CLK 835 |

Qy 655 - - - - - H P E S S K S H S R A A S A L W N M L E P V G Q K T T R H K F E N K I N I T C P N Q V O P F L R T 703

Db 836 FCTAAISYSLCKFSSQSRTDLCSCLSPGLIK-----KEQFLMRLFSEARQPLSER 886

QY 760 DIPDVTTQ-----YRGlySAGGDSIENV-TLPINLREERNLIGAVGTGQ 807

Db 945 ALSDTERQDFHQWAIEHFLPESSASGGCGDGLQACTILIVNALMDFHQSRSY-----D 999

1000 HSEN----SULVFGKTCNEDILSKQEMVAJUE----LQOD---LIVPLGHP---- 104
866 CTDSSLYSAANFV-DHLRNALDDVL----HREV---PRVVLNLDFLNPTIMR 909

Db 1043 ---SOERHFLFEIIRRRLQALTSGWSVAASLQRORELIMYKRILRL--- 108

QY 910 QVFLGNPDKCPVQASVTCNCVLTURENSQELARLEAFSRAYRSSRELVGSGRYDQED 969

8Y DB 1086 -----PSSVLCGS--SFQAEOPITARCEOFFHLVNSEMRNFCSHGGALTQD- 112
 970 FSVLLOPFFONIOLQPLVADGLP--TSFFAPPCIHPNOKPHISOLARL-WTNLLEPIGSK 103

Db 1130 ---; ITAHHFFRLNACLRSRDPPLSMVDFILAKC---QTKCFLILTSALUWMPSLPEV-- 118

OY
1027 TETLDLRAEMPTCPTONEPFLRTPRNSNYTPIKPAIE--NWGSDFLCTE----- 107

Db 1181 -----LLCRWRRHQCSPPLRE-----LQKQSGRQFASPLSPEAASPAPNP 122
Qv 1076 WKANSVPUVTYHORPANTK-----WVAAGNSTTAWAPPN 1111

*1
Db
1224 WLSAAALHFAIQQQREENIRKQLKKUDCEREBEPLLVLFLFFSSLMGLISSLHTS---- NS 127

OY 1115 SSDLPTSW-----RGLSW 1127 :::: :::: ::

Db 1278 TTDLPKAFHVCAILECLEKRKISW 1302

RESULT 3
US-09-589-567-2

; Sequence 2, Application US/09589567
; Patent No. 6479730
; GENERAL INFORMATION.

APPLICANT: Mahajan, Pramod B.
APPLICANT: Shi, Jinrui

; TITLE OF INVENTION: Maize DNA Ligase II Orthologue and Uses
; TITLE OF INVENTION: Thereof
FILE REFERENCE: 1125

CURRENT APPLICATION NUMBER: US/09/589,567
CURRENT FILING DATE: 2000-05-07
DRAFTING ATTORNEY NUMBER: TIG 50/145 911

PRIOR APPLICATION NUMBER: US 07/145,211
PRIOR FILING DATE: 1999-07-27
NUMBER OF SEQ ID NOS: 3

; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 2
; LENGTH: 1218
; TYPE: PRT

RESULT 4
US-09-180-439-B
Sequence 8, Application US/09180439
; Patent No. 6225532
; GENERAL INFORMATION:
; APPLICANT: Hatzixanthis, Kostas
; APPLICANT: Jones, David A
; APPLICANT: Jones, Jonathan DG
TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof
FILE REFERENCE: 620 - 53
CURRENT APPLICATION NUMBER: US/09/180,439
CURRENT FILING DATE: 1998-12-06
; EARLIER APPLICATION NUMBER: PCT/GG97/01249
; EARLIER FILING DATE: 1997-05-08
; EARLIER APPLICATION NUMBER: GB 9609681.3
; EARLIER FILING DATE: 1996-05-09
; EARLIER APPLICATION NUMBER: GB 96119924.5
; EARLIER FILING DATE: 1996-09-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 1016
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
; US-09-180-439-8

Query Match
Best Local Similarity 19.6%; Score 127.5; DB 3; Length 1016;
Matches 207; Conservative 134; Mismatches 380; Indels 333; Gaps 55;

Qy 40 TAKNSPPCNPNIKGUNPKSKVHSLSKSDIKVAAGNLT-----ERPD 84
Db 85 TYAPPFSSLPLLENLDLSNNNTGIPPE-----IGNLTLYVDLNTNQDISGTIPQ 138
Qy 85 PGTCDELEKDWTE-----RPQQV-CMGWMTLSDIRTRSPSPVMPVCGTKRVIP 134
Db 139 -IGSLAKLQITRIFNNHNGFPEEGYLRSLTKLSGGINFLSGSIPASLGNN----- 190
Qy 135 HDGAEDLWTOAQELVRNME-----NLODQFOF-----DWKLINVFESNASOCYLC 180
Db 191 -TWSFLFLYENOLSGSIPPEBIGYRLSLTKSLDINFSGSIPASLGNL-MNLFLYL- 246
Qy 181 PSAQONGQLAAGCYDELGMVTL-----DLYQEVRAFYVLVLDUSEAVARSQHGT 230
Db 247 -YNNOLSSIPPEBIGYRLSLTKSLGINFSGSIPASLGNNLNSRDLJNKLGS 302
Qy 231 WLSPAPEPONCSBTTRLAKVNOMSYOBANNSLLASSRYSEQESFTVFQOPFFY-BTT 288
Db 303 ---IPEEIGYRLSLT-----YLDGENAANGSI PASLGNLNLLFMLYNNLGS 348

RESULT 5
US-09-252-991A-25753
Sequence 25753, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIORITY APPLICATION NUMBER: US 60/074,788
; PRIORITY FILING DATE: 1998-02-18
; PRIORITY APPLICATION NUMBER: US 60/094,190
; PRIORITY FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25753
; LENGTH: 2636
; TYPE: PRT

Qy 289 PSLHSEDERLQDSTTLAWHLWNRMMEPAGEKDEPLSVRKGPRMKPCSQESPYLPYSRNSN 348
Db 349 GSPEEIGYRLSITL-----DLEGAVNLGSI-----PA-----SIGLN 383
Qy 349 YLTJLQKEDKLE-----VREGAEIREPDKDSSTDVTSPVHLK----- 387
Db 384 NLSDLVLYNNKLSGSIPEEIGYRLSITLVDLGENALNGSIPASLGNLNLFMLYLNOL 443
Qy 388 ---PADINVIGLGSITAGNGA-GSTP---GVNLVLTQY---RGLSMSVGGENIG 435
Db 922 GNEL-----MSFOELSSRERGNKHSSISI-----ENIKUDICVFDFDIMFDG 964
Qy 990 -----LPDTSFAPDCITHPNOKFHSOLARALWTNMLEPLASKTETDLREMPI---- 1038
Db 965 ERILDCPQRQRKCYTHDFHEKGYF ELAQNLNEACEASPDNSTL--ARMDTFFRN 1020
Qy 1039 TCPTONBPFPLTRPRNSNTYPKPAIENWGSDFLTCTEWKASNSVPTSVHOLRPADKV-- 1096
Db 1021 ACOSCEGIMLKTLVDAGYSASKRCDSW-----LKVR 1054
Qy 1097 -YAAIGSL-TTAVGA 1110
Db 1055 DVVQGLGDBSLDLYVPIGA 1071

RESULT 4
US-09-180-439-B
Sequence 8, Application US/09180439
; Patent No. 6225532
; GENERAL INFORMATION:
; APPLICANT: Hatzixanthis, Kostas
; APPLICANT: Jones, David A
; APPLICANT: Jones, Jonathan DG
TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof
FILE REFERENCE: 620 - 53
CURRENT APPLICATION NUMBER: US/09/180,439
CURRENT FILING DATE: 1998-12-06
; EARLIER APPLICATION NUMBER: PCT/GG97/01249
; EARLIER FILING DATE: 1997-05-08
; EARLIER APPLICATION NUMBER: GB 9609681.3
; EARLIER FILING DATE: 1996-05-09
; EARLIER APPLICATION NUMBER: GB 96119924.5
; EARLIER FILING DATE: 1996-09-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 1016
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
; US-09-180-439-8

Query Match
Best Local Similarity 19.6%; Score 127.5; DB 3; Length 1016;
Matches 207; Conservative 134; Mismatches 380; Indels 333; Gaps 55;

Qy 40 TAKNSPPCNPNIKGUNPKSKVHSLSKSDIKVAAGNLT-----ERPD 84
Db 85 TYAPPFSSLPLLENLDLSNNNTGIPPE-----IGNLTLYVDLNTNQDISGTIPQ 138
Qy 85 PGTCDELEKDWTE-----RPQQV-CMGWMTLSDIRTRSPSPVMPVCGTKRVIP 134
Db 139 -IGSLAKLQITRIFNNHNGFPEEGYLRSLTKLSGGINFLSGSIPASLGNN----- 190
Qy 135 HDGAEDLWTOAQELVRNME-----NLODQFOF-----DWKLINVFESNASOCYLC 180
Db 191 -TWSFLFLYENOLSGSIPPEBIGYRLSLTKSLDINFSGSIPASLGNL-MNLFLYL- 246
Qy 181 PSAQONGQLAAGCYDELGMVTL-----DLYQEVRAFYVLVLDUSEAVARSQHGT 230
Db 247 -YNNOLSSIPPEBIGYRLSLTKSLGINFSGSIPASLGNNLNSRDLJNKLGS 302
Qy 231 WLSPAPEPONCSBTTRLAKVNOMSYOBANNSLLASSRYSEQESFTVFQOPFFY-BTT 288
Db 303 ---IPEEIGYRLSLT-----YLDGENAANGSI PASLGNLNLLFMLYNNLGS 348

RESULT 5
US-09-252-991A-25753
Sequence 25753, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIORITY APPLICATION NUMBER: US 60/074,788
; PRIORITY FILING DATE: 1998-02-18
; PRIORITY APPLICATION NUMBER: US 60/094,190
; PRIORITY FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25753
; LENGTH: 2636
; TYPE: PRT

; ORGANISM: *Pseudomonas aeruginosa*
US-09-252-991A-2573

Query Match 1.6%; Score 123.5; DB 4; Length 263;
Best Local Similarity 21.5%; Pred: No. 0, 2;
Matches 154; Conservative 83; Mismatches 264; Indels 215; Gaps 39;

Qy 814 FVNQAVPGAK-----AEDLMSQVTLQRMKMDHRVNFHE---DWKV--I 853
Db 1121 FEQQPIPRQHNNQLSLLKPREALNKALEALQALYEH-HDALKRKFETGTMWAEHA 1179

Qy 854 TFLIGGSDICPDCTDSMILYSAANFVDIHLRNALDVHLREVPRLVNLNDFLNNTIMROYL 913
Db 1180 EATIGGA-----LILWRAEVD--RQALESLCBESORSL---DIAOCPILLRSLLV 1223

Qy 914 GPPDKCIVQOASVLTNCVLTILRENSQBLARLAFSRAVSSR---ELVG-----S 961
Db 1224 --DMADGGQRILLVHLILVNGVSWRL-LIDLQRQVOSRGEARPLPGTISPEKAWA 1279

Qy 962 GRYDTQ---EDPSVUOPLLFFQNTOLPVLADGPDTSFAPDCIHP---NOQP---HSQ 1010
Db 1280 GRVSEHARGESKQAKAQLO-FWRG---LLEGPAEL---PCVHPQAGALEQRAFSVQSR 1329

Qy 1011 LARALWNTMLE--PLGSKTEID--LRAEMPPTCPTONEPEPRTPRNSNYTPKAIEN 1066
Db 1330 FLRSLLTERLKKQAPAAYTQVNDLILTALARVC-----RNWSGASSL-VOLEG 1377

Qy 1067 WGSDFLCTEWKASNSVP-TSHQLRADIKVAALGDSLTT---AVGCARPNSSLPLTS 1121
Db 1378 HGRBELFADIDISRTVQWFTS--LFPVRLSPVADIGEBSLKAIKEQRAIPD----- 1426

Qy 1122 WREGLSWS---IGGDGNLETHTTLPLNITKFNYPYLGPSTSTWEGTAGLNVAEGARAD 1177
Db 1427 -KGIGYGLRYLJAGEESARVLAGLPQARITFN-YLGOPDAOPDEMALLDAGESAGEM 1483

Qy 1178 MPAQAMWLVKRNNSPIN-----LRKDWNKLVTLEFIGVNLCHYCENPEAHAYQ 1230
Db 1484 DPGAPLD-----NWLSLNGRVFDGELSIDNSWFFSSQMPG-----EDQVRLLADDYA 1529

Qy 1231 HQQQLALISEB-----LPAFVNNEVMELASLYQOGOKCA 1268
Db 1530 ETALVUDVCCDSPRRGATPSDPLAGLDQARLDALPVALYEEDIVPLSPMQQ----- 1583

Qy 1269 MLI-----AQONCTCLRHSSOSLEKQELKUNWNL---OQGISSFSYWHYOTBDDFA 1318
Db 1584 MIFHSLYEQASSDYINQMRVUDVSGLDIPRFRA-WOSALDRHAILRSGFWAOGELQOPLQ 1642

Qy 1319 VVVOQPEFQNTLPLNERGDTDLTFESDCFHSDRGAEMAALMNMLLEPGRKTSNN 1378
Db 1643 IVYR-----ORQDPAEEDLSQANRDAALLA-----AAERERG 1678

Qy 1379 FTHSRAKL-----KCFSPESPYLYTURN-----SRILPQAEPEVL 1416
Db 1679 FELQARPLRLILVKTEGEHHLYTTHHILLDGWSNAQLAESVLSSESYAGSPEQL 1734

RESULT 6 US-09-107-532A-5646

Query Match 1.6%; Score 122; DB 4; Length 855;
Best Local Similarity 19.4%; Pred: No. 0, 29;
Matches 134; Conservative 88; Mismatches 237; Indels 230; Gaps 36;

Qy 357 QDKLEYREGAEIIRCPDKD--ISDTPVPSVHLKPADINVIGA---LGDSLUTAGNGAST 410
Db 118 QBAIVLNSGMOKBKIPEAKLVKGDVIVLRAGDVWPADIRLIERHDLIVEESLTGESEASE 177

Qy 411 PGAVLVDLTOVGLSWSVGGDENIGTVTLANLIREBPNPSLJKGFSVGTGKETSPNAPFLNQ 470
Db 178 KSHA--VLTEEB---SIGQKQMGKSGTCVQ---SGSAVGUVVETGNTIEKG-INQ 225

Qy 471 AVAGGRABDLPVQARRVILDKMN-----DTRHFQEDWKITLFIGNDLC 516
Db 226 AQSVEQQTLP-IKKINRNLNQIQFOGILCISLFLVIFTFRYGMWHIL----- 274

Qy 517 DFCNDLHVYSQNFTNIGKALDILHAEVPAFVNULVLEIVNRELQEKVV--CPR 574
Db 275 -----LSARTALVUWSMVPBGLPAVLTIMLISVGHEMAKEKAIKGPLS 317

Qy 575 M----IIRSLC-----PCVLFKDDDNSTELATLIEFNKKCQKT 608
Db 318 VETLGSMVTCSDKTMITKREMTVUDVAAKEPACVLSIMCNGQELKT-----KEEQKT 371

Qy 609 HOLIESGRYDTRBDFTVVWVQPFENYDMPKTSSEGPDNSFF-----

Db 372 ENL--SG--NPTEVALRIRHETANPLKKTEAKIPPFSEYKMATMHAEKGAVIVKG 426

Qy 650 APPCFHRSKSHSRAASALWNMLEDVGQKTR-----HKFENKINTCENQVQPLRT 703
Db 427 APEVL-FAKSTLSSABQEAWSQTAEPARKGQRLGVFKVDSKOBLTHTLTO---LT 482

Qy 704 YKSMQCHGWTLPCCRDRAPSALHPTSVHLRADIQVVAALGDSLTAGNGISKPEDPLD 763
Db 483 FAGIA---GLDPPKESAVKA-----VKECQOAGISVYKMTGHD-----KDTAK 523

Qy 764 VVTOYGLSYSA---GGDGSI---ENVTL-----NIREFNRNLTY 801
Db 524 AIAEQIGLKHTAKVLSGIDDLMSDBELLIQQVPIVDPARTPEHKURIVKALOKN--GE 581

MEDIUM TYPE: CD/DVD ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107-532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085, 598
FILING DATE: 14 May 1998
REFERENCE NUMBER: 60/051571

ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Dencke

REGISTRATION NUMBER: 40, 489
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 5646:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids

TYPE: amino acid
TOPOLOGY: linear
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1... 855

SEQUENCE DESCRIPTION: SEQ ID NO: 5646:
US-09-107-532A-5646

Query Match 1.6%; Score 122; DB 4; Length 855;
Best Local Similarity 19.4%; Pred: No. 0, 29;
Matches 134; Conservative 88; Mismatches 237; Indels 230; Gaps 36;
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
NUMBER OF INVENTIONS: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:

QY 802 AVG-TGD-ANDTNAFLNOAV---PGAKAEDLMSQVQTLMQKODHRVNFHDWKLKVITVL 856
 Db 582 RKGWTGDCVNDA-PALKSDPVGJAMGIKSEVSQVQAADMV--LGDD--NEHTIAKAV-- 633
 QY 857 IGGSDLCDYCTPSNLYSANFVDFHLRNALDVHREVPRLVNL---VDFLNPTIMRO--- 910
 Db 634 -----KEGRRLDNLQKTNFFLPPTALAQLI 660
 QY 911 ---VFLGNPDKCPVQQASVL-CNCVULTL 934
 Db 661 LTMVALMLNRP-LPLSPVQILWNVTTI 687

RESULT 7
 US-09-353-585-2
 ; Sequence 2, Application US/09353585
 ; Patent No. 628865

GENERAL INFORMATION:

APPLICANT: Dixon, Mark S
 Jones, Jonathan DG
 Jones, David A

TITLE OF INVENTION: Plant pathogen resistance genes and uses
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Nixon & Vanderhye PC
 STREET: 8th Floor, 1100 No. 6287/865th Glebe Road
 CITY: Arlington
 COUNTRY: United States of America
 ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/353, 585
 FILING DATE: 15-JUL-1999
 CLASSIFICATION: C12N 15/29, 15/62, A01H 5/00, A01N 65/00, C12Q 1/68

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/930, 277
 FILING DATE: 27-OCT-1997

APPLICATION NUMBER: PCT/GB96/00785
 FILING DATE: 01-APR-1996
 APPLICATION NUMBER: GB 9506658.5

FILING DATE: 31-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Ms. Marty J. Wilson
 REGISTRATION NUMBER: 32, 955
 REFERENCE/DOCKET NUMBER: 620-69

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4100
 TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 112 amino acids
 TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE: ORGANISM: Tomato
 STRAIN: Cf2
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-353-585-2

Query Match 1.6%; Score 121.5; DB 3; Length 1112;
 Best Local Similarity 18.2%; Prcd. No.: 0.055; Mismatches 356; Indels 387; Gaps 48;

QY 53 LGVNMPSPSV-----HSLK-----PSDIFVAIGNILEIPPDGTGDEK 92
 Db 174 LGINFLSSIPASPGVGNLNLFLYLYNNQLSISIPERISVLSLTEDLSDNALNGSI-- 231
 QY 93 QDPTTERPQVCMGVMTVLSIDIRY--FSPSPVMPVCHTGKRVIPHQGAED----- 140
 Db 232 -----PASLGNMNLNSFLFLYGNOLGSISPEECYL-RSLTYLDSENALNGSTPAS 282
 QY 141 -----LWIAQAEI-----VRNNE----- 154
 Db 283 LGINLNLSPFLFLYGNOLGSISPEBEGYRLSLVNLGLUSENALNGSTPASLGNKLNLSNL 342

QY 155 -NIQOLDFOFDWKLININVFFNSASOCYLCPSAQONGLAAGVDELMGVIDLQO----- 205
 Db 343 VNLQLSGSSIPASLGNL--NNLMLY---TNNQLSGSSIPASLGNINNLMLYLYNNQL 395

QY 206 --EVPRAVNLVLDSEVEAVRSOYHGTWLSPAPEPCNCSEETRLAKVUMONSYEARNS 263
 Db 396 SGSSIPASLGNLNLNSFLFLYGNOLGSISPEECYL-----SPEEIGYLSITYLDLSNNS 445

QY 264 LLASS-----RSEQESPFTVURQPFYETPSLSDPRQDSTTAWHLMRM 312
 Db 446 FIASPGMNSNLAFLEFLENQASSVPEBEGYRLSLVNLDSENAALNGSTPASLGNKLNLSNL 505

QY 313 ME PAGEKDEPLSVKHGPCKPSQESPYL-----FSYRNSYNTLQ 354
 Db 506 SR-----LNUVNQNLGSPEEIGYRLSLVNLDSENAALNGSTPASLGNKLNLSNL 557

QY 355 KRODKLEVREGABIRCDDKDPSDTVPSPVHRLKPADINVIGALGDSHTAGNAGSTRG-- 412
 Db 558 LVNNQL-----SGSIPPEBEGYR--SLNDGLSENALN-----GSIPASL 595

QY 413 --NVLDVLTQXRG-LSVSGEDENITGVTUANIREFNPLKGFSVGTGKETSPTAFL 468
 Db 596 GNANNLNLNLNTYNNQNSI--PEELCYLSSI-TYNSGNSINGL-----IPASA 644

QY 469 NQAVAGGERADELVEQARRELNDLJKNDTRHFQDWKITITLFIG--GNDLCDFCNDLYHS 526
 Db 645 NM-----RNQALIJLNDNL-----IGEIPSSVCNTSLEYLM 678

QY 527 PQN-----FTDNIGKAL-----DILHAVERA 548

Db 679 PRNLLKGKVPOCLGNISNLQVLSMSSNSFSGELPLSSISNLISLQLDFFGRNNELEGIAFCQ 738

QY 549 FVNLVTVLVIWIRE----LYQEKKYCPMIL-----RSLCC-----VTK 586
 Db 739 FGMI-ISSEVFDMQNNKLSGTPTNFSIGCISLISLNHGNELEDETPRSLDNCKQQLVD 797

QY 587 FDDNSTE-----LATLIEFN--KKFOEKTHOLIESGRYDIREDFTVVQPFENVMFK 638
 Db 798 LGINQLNDTFPMWLGTLPELVRULTSNKLHOPRIRSRAE-----IMFPDLRUDSLR 850

QY 639 T--SEGIDPNSSFAPDCFHFSKSKSHSRASALWNMPLPVGOK----- 679

Db 851 NAFSQDLPTSLF-----EHLKGMRVTDKTMEBPSYESYYDDSSVVTKGLEI 899

QY 680 -----TTRHKPENKINTICPKVQO-PLRTKNSMCGHGTWLPCRDRAPSILHP 727

Db 900 VRILSLYTVIDLSKNCBEGHSVGLDIAIRLUNSHANALOGY-----IPSSIGS 950

QY 728 TSYHALRADIQVVAALGDSLTAGNGIGSKPDLPDT-TQYRGLS--YSAGGDGSLEN 784
 Db 951 LSI--LESQD-----LSPNQSGSEIPQOLASITFLEPLNLNSLYQCGCIPQQRQP 998

QY 785 TTLPNLTLEBFNRNLITGYWAGTGDANDTNAFLNOAVPGKAEDLMSQVTLWQMKDHRV 844

Db 999 RTPESNVSEGNDGLRGYPVSKGGKDPVSEKVNVSALEDQESNSR----- 1045

QY 845 NFHDWKVITVUG-GSDLC-----DYCTDSNLISANFVDFHLRNALDVHREVR 894
 Db 1046 -FNUFWK--AAUWYGGSLCIGISMIVLISITGNCRLMARIEKLEKIKIMORRKCR 1100

RESULT 8
US-07-741-940-7
; Sequence 7, Application US/07741940
; Patent No. S352775
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; APPLICANT: ANAND, RAKESH
; APPLICANT: CARLSON, MARY
; APPLICANT: GRODEN, JOANNA
; APPLICANT: HEDGE, PHILIP J.
; APPLICANT: JOSLIN, GROFF
; APPLICANT: KINZLER, KENNETH
; APPLICANT: MARKHAM, ALEXANDER F.
; APPLICANT: NAKAMURA, YUSUKE
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/741,940
; FILING DATE: 19930109
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.035574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-8299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2842 amino acids
; TYPE: AMINO ACID
; STRANDBEDNESS: Single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: APC
; US-07-741-940-7

Query Match 1.6%; Score 120.5; DB 1; Length 2842;
Best Local Similarity 19.8%; Pred. No. 046; Mismatches 328; Indels 311; Gaps 47;
Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

Qy 34 GQWQPETIKNSTPPCNPENPKLGVNMPSKSVHSIKPSDKIFVVAALGNLELPPDGTCGDEK- 92
Db 1920 GQPKPQILOKQSTTP-----OSSKDI-----PDGAATDEKL 1950
Qy 93 QWTERPQQVCM---GIVMTLVDIIR-----YFSPSVPM----- 123
Db 1951 QNPAENTPVCFSHNSSSSLSDQEBNNKENKEPIKETEPDPSQGSKPSKPGASVARKS 2010
Qy 124 ---PVC---HTGKRVIPHGAEDWIAQABLV-----RNM 152
Db 2011 FHVETPVCFSNSLSSLSDSEDDL---LQCISSAMPKKKPSRUKGDVKHSFRM 2067
Qy 153 ---KENQLDFO-----FDKLI---NPFPSNLSQ---CYCLCPRAQ 185

RESULT 9
US-08-289-548A-7
; Sequence 7, Application US/08289548A
; Patent No. 5689212
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; APPLICANT: ANAND, RAKESH
; APPLICANT: CARLSON, MARY
; APPLICANT: GRODEN, JOANNA
; APPLICANT: HEDGE, PHILIP J.
; APPLICANT: JOSLIN, GROFF
; APPLICANT: KINZLER, KENNETH
; APPLICANT: MARKHAM, ALEXANDER F.
; APPLICANT: NAKAMURA, YUSUKE
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, LTD
; STREET: 1001 G Street, NW
; CITY: Washington

Db 2068 GGILGEDILTLDKIOPPDSERGLSPDSENFDMKAIOBGANSTVSSUHQAAAACLSRQ- 2126
Qy 186 NGIAAGGDPDELMGVLDVYDQVEPRAFTVNLVSEEVASVROVHGTVLSPAPPCNCSET 245
Db 2127 --ASSDSDSILSILKSGISLGSP--FHLPDQEKEPFTSNK-----GPRLKPGKS 2173
Qy 246 TRLAKVVMQWSY----OBAMNLLASSRSYSEOBSTVVFQOPFYETPSLHSEDPRQD 300
Db 2174 TLETKKKESSESKIGKGKVVNLITCKVRSNEISEGOMKQ-----LOANMPSTSR 2225
Qy 301 STTLAWHJNRMMEPAGEKDEPSVKGKRPMKCPSQSOPSPYLFSYRNLYNLQKQDQL 360
Db 2226 GRIMI-HIPG--VRNSSSSTSVPS-KQGPPLKTPASKSP--SEGQTTATSPRAKPSWS 2279
Qy 361 EV----REGAETRCPDKPS----DTPVTSVTRLKRDADINTIGALGDSLTAQNGAGSTP 411
Db 2280 ELSPVARQTSOIGGSSKAPSRSQRSDSTPSRPAQ-QPLSPRIQSPGRNISPGNRGNSPP 2338
Qy 412 GNVLVDVLTQYRGLSWVSGDENIGTVTIANLREFPSLKG-SVETGK--ETSPNAPL 468
Db 2339 NK-----LSQUPRTSSPSTASTKSSGKGMSYTSPEQM 2372
Qy 469 NOAVAGGRGAELPVQARRLVNDKNTTRHFQD-WKITLTFTGGNDICDFCDNLVHSP 527
Db 2373 SQ-----QNLTKQ-----TGLSKMASSIPRSSEASAKQNLQNMNG---ANKKVELSR 2417
Qy 528 QNPTDNTICKALDITLHAEPVRAFPVNLYVLTETVNLREYQEKVYCPRMILRSLCPVULKP 587
Db 2418 MESTKSSSESDE--RSERP-----VLRQSTPIKE-----APSPTRL----- 2452
Qy 588 DDNSTELLATLIEPNKKFQEKTHOLIESGRYDTRBDFTVVVQOPPFENDMPKTSBGLUDNS 647
Db 2453 -----RKLESZASFESLSPSSRSPASPITRSQTPTVPSLUD-- 2489
Qy 648 FFADPCHFSSKSHSRAASALNNM--LEPV----GQKTTTRHKFEENKINTCPCNQWQ 699
Db 2490 -----MSLTHSSVYAGGWRKLPPNLSPTBYNDRPAKHDIA-RSHSESPRL-P 2539
Qy 700 PLRTYKNSMQHGTVLWCRDRPALSPTVHMLRPADIQVIALGSLTACNGIGSKPD 759
Db 2540 INRS-----GTWKRBHSHSSSLPRVSTWARTESSSSILSASSESEK----AKSE 2586
Qy 760 DLPDVTFQYRGLSYAACGDSLLENVTLPNLR----EPN-RNLTYAVGCDANPTNA 813
Db 2587 DEKRVNS-----ISTKTKSOKENOUSAGTKRKIKEFESPNTSQTSVSGATN---- 2635
Qy 814 FLNQAVPGKAEDILMSQVOTLMQRKD 840
Db 2636 -----GAESKTLIYQMAPAVSKTE 2655

STATE: D.C.
 COUNTRY: USA
 ZIP: 20001-4598

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/289,548A
 FILING DATE: 12-AUG-1994
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Kagan, Sarah A.
 REGISTRATION NUMBER: 32,141
 REFERENCE/DOCKET NUMBER: 1107.46943

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-508-9100
 TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
 LENGTH: 2842 amino acids
 TYPE: amino acid
 STRANDEDNESS: single

TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 IMMEDIATE SOURCE:
 CLONE: APC

US-08-289-548A-7

Query Match Best local Similarity 19.8%; Pred. No. 0.46; Mismatches 328; Indels 311; Gaps 47; Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

Qy 34 QLWMPETIJKNSPPCPNPNKLGIVNNPMSKVKHSLKPSDIDKREVALGNLEIPPDGCTGDJLK- 92
 Db 1920 GQPKPILQKOSTFP-----OSSKDI-----PDRGAATDBKL 1950

Qy 93 QWTERPOQVCM---GWMVTLSDLIR-----YFSPSPVPM----- 123
 Db 1951 QNFAILENTPVCFHSNSSLSSLDQENNNKEMPIKUTEPDPSQGPFSKPKQASGYARKS 2010

Qy 124 -----PVC----HNGKRVIPHAGEDLMQAOILV----- 152
 Db 2011 FHVEDTPVCFSRNSLSSLSISDSEDDL--LQBCISSAMPKCKPSRLKGDIKHSFRM 2067

Qy 153 ---KENQDFQ-----FOMKLI---NIVFSNASQ-CYCLCPAQO 185
 Db 2068 GGTIGLEDLTDLKDIQRPDSEHGLSPDSENFDWKAIORGANSITVSSLHQAAAACUQR- 2126

Qy 186 NGLAAGGYDULMGVLQDQEVNPRAFVNLDISEVAEVSRQHGTLWLPAPCPCNCSEBT 245
 Db 2127 --ASSDSISLISKSGISLGSF-FHTPDQEKPFTSNK-----GPRILKPGKSK 2173

Qy 246 TRIAKUVQWWSY---QBANWISLASSRYSBOSBFTVWFQPFYETPPLSISDPRQLD 300
 Db 2174 TLLTKKIESESKGIGKGRKVYSLTGVRNSBSISGQMKP-----LQANMPISR 2225

Qy 301 SITLAWHWHWRNRMMEPAGEKDEPLPSVKGIGRPMKPSQESPPYLTSYRNNSVLTQKQDKL 360
 Db 2226 GRIMI-HIFG--VTRNSSSSTPSV-KKGPPLKTPASKPS--SBRGQATATSPRGAKPSVKS 2279

Qy 361 EV----REGAEATRCPDSSPS----DTVPTSVRLKDADINTVAGALGDSLTAQNGAGSTP 411
 Db 2280 ELSPVARTSQIGGSSKAPSRSRSDSPSRPQ-QPLSRPQISOPGRNISISORNGISPP 2338

Qy 412 GNVLDLVLTQYRGLSWSVCGDENINGTVTMLANTIREFENSLKRP-SVGRGK--ETSPN AFL 468
 Db 2339 NK-----LSQLPRTSSPSTAATKSSGSGKMSYSPGRQM 2372

469 NAVAGGARAEDIPLVQARLVLDMNDTRIHPQED-WKITUFIGGNDLCDFCNDLVHYP SP 527

RESULT 10
 Sequence 7, Application US/08452654
 Patent No. 5691454

GENERAL INFORMATION:
 APPLICANT: ALBERTSEN, HANS
 APPLICANT: ANAND, RAKESH
 APPLICANT: CARISON, MARY
 APPLICANT: GRODEN, JOANNA
 APPLICANT: HEDGE, PHILIP J.
 APPLICANT: JOSLIN, GROFF
 APPLICANT: KINZLER, KENNETH
 APPLICANT: MARKHAM, ALEXANDER F.
 APPLICANT: NAKAMURA, YUSUKE
 APPLICANT: THIERVERIS, ANDREW
 TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
 NUMBER OF SEQUENCES: 94
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Banner, Birch, McKie & Beckett
 STREET: 1001 G Street, NW
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20001-4598

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/452,654
 FILING DATE: 25-MAY-1995
 CLASSIFICATION: 536

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/741,940
 FILING DATE: 08-AUG-1991

ATTORNEY/AGENT INFORMATION:
 NAME: Kagan, Sarah A.
 REGISTRATION NUMBER: 32,141
 REFERENCE/DOCKET NUMBER: 1107.035574
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-508-9100
 TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
 LENGTH: 2842 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE: Homo sapiens
 IMMEDIATE SOURCE: APC
 CLONE: APC
 US-08-452-654-7

Query Match 1.6%; Score 120.5; DB 1; Length 2842;
 Best Local Similarity 19.8%; Pred. No. 0.46; Length 2842;
 Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

QY 34 GOLWPETLKNSSPPCPNPKLGVNMPSKSVHSLKPSDIKFVAIGNLEIPPDGTGDLK- 92
 Db 1920 GQPKPILQKQSTFP-----QSSKDI----- 92
 QY 93 QMTERPPOQVCM--GWTMVLSPDR----- 123
 Db 1951 QNFAIAGENTPVCFSHNSLSSLSDIDQENNNKENEPIKETEPPDSQGPSPSKPAGSYAPKS 2010
 Qy 124 -----PVC--HTGKRVIPHOGAEDIWIAQDLY----- 152
 2011 FHVEDTPVCFSRNSSLSSLDSEDDL---LOBCISSAMPKKKPSRLKGDNKEHKSFRM 2067
 Qy 153 ---KENIQQLDFQ----- 185
 Db 2068 GGTGLEDITDIDKIQRDSENLGSQDSENFKAIQEGANSIVSSLUQAAACLSRQ- 2126
 Qy 186 NGLAAGGYDELGVLDYLQEQEVRRAFLVNLSEVAEVSRQTHGTWLSPARPCNCSEBT 245
 Db 2127 --ASSSDSISLKLKGSLGLSP--FHLTPDQEKKPFTSNK-----GPRILKPGEKS 2173
 Qy 246 TRIAKVWQWSY----QEAWSLLASRSYSESFTWVQDFFYETPSLHSEDPRLQ 300
 Db 2174 TBTKKLESSESKGIGKKVYKSLITGKVRSNEBISGQMKP-----LOANMPISR 2225
 Qy 301 SITTLAWHINWRNRMPEPAGEKDEPLSVKHGRPMKCPQSOSPYLSYRNNSYLTQPKDQ 360
 Db 2226 GRIMI-HIPG--VRNNSSTSPIs-KCOPPLKIPKASKP--SEQQTATTSRPGAKSVK 2279
 Qy 361 EV----REGAEETRCPDFKDP-----DRTWPTSWRLKRADINTVIGALCDSLTGNGAGSTP 411
 Db 2280 EISPVAROTSQLGGSSKAPSRSSRSDSTPSRPAQ-QPLSRPQISPGRNNSISGRNGISPP 2338
 Qy 412 GNVLDVLTQYRGISMSWVGDDENIGTVTIANLREFFPLSKP-SVGTGK--ETSPMAFL 468
 Db 2339 NK----- 2372
 Qy 469 NQAVAGGRAEDLIVQARRLVDLMKNDTRIFCQD-WKITLRTGGNDLICDFCDLVHNSP 527
 Db 2373 SQ-----QNIJKQ---TGLSKNASSIPRESASKLNQMNNG---ANKKBLSR 2417
 Qy 528 QNFTDNCALDITHAEVPRAFVNLTVEIINRELVQEKVYCPRMILSSLCPCVLF 587
 Db 2418 MISTKSSGSESD--RSERP-----VLVRQSTFIKE-----APSPTLR----- 2452
 Qy 588 DNSTELATLIEFNKKQBKTHOLIESCRYDREDFTVWQOFENNMDPKTSEGUDNS 647
 Db 2453 -----RKLEESASFESLSPSSPSSRASPRTSOAQTPVLSPLD-- 2489
 Qy 648 PRAPDCFHFSKSHSRAASALMMN--LEPV----GQKTRHKFENKINTCPNOVQP 699
 Db 2490 -----MSLTHSSVQAGGRKLPPLSPTEYNGRPAKGHDIA-RSHSEPSKL-P 2539
 Qy 700 FRTYKNSMOGHETMLPCRDRAPSLHPTSVAIRPADIQVVALGDSLTAENGICSKPD 759
 Db 2540 INRS-----GTWKREBSKSSLPRVSTWRRTGSSSIIASSESSEK----AKSE 2586
 Qy 760 DLPDVTTQYRGISYSAGGDGSLLENVTULPNLUR----EFN-RNLTYGAVGCGDANDTNA 813

RESULT 11
 US-08-449-731-7
 Sequence 7, Application US/08449731
 Patent No. 6413727

GENERAL INFORMATION:

APPLICANT: ALBERTSEN, HANS
 ANAND, RAKESH
 CARLSON, MARY
 GRODEN, JOANNA
 HEDGE, PHILLIP J.
 JOSLIN, GEOFF
 KINZLER, KENNETH
 MARKHAM, ALEXANDER F.
 NAKAMURA, YUSUKE
 THLIVERIS, ANDREW

CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20001-4598

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.
 NAME: RAGAN, SARAH A.
 ATTORNEY/AGENT INFORMATION:
 REFERENCE NUMBER: 32-141
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-508-9100
 TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
 LENGTH: 2842 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE: Homo sapiens
 IMMEDIATE SOURCE: APC

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-08-449-731-7

Query Match 1.6%; Score 120.5; DB 4; Length 2842;
 Best Local Similarity 19.8%; Pred. No. 0.46; Length 2842;
 Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

QY 34 GOLWPETLKNSSPPCPNPKLGVNMPSKSVHSLKPSDIKFVAIGNLEIPPDGTGDLK- 92
 Db 1920 GQPKPILQKQSTFP-----QSSKDI----- 92

93 QDNTTERPOQVCM--GWTMVLSPDR-----YFSPSVPM----- 123
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 1951 QNFAIENTPVCFSHNSLSSLSIDQENNNKNEPIKEPPDSQGERSKPKQASGAPKS 2010
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2011 FHVEDTPVCFSRNSLSSLSIDSEDDL--LQCISAMPKKPSRLKGDNKHSRNM 2067
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 153 ---PVC---HTGKRVIPHDGAEDLWIAQBLV-----RNM 152
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2068 GGILGEDLTLDLKDQIRDSEHGLSPDSENFDWKAIORGANISIVSLHQAAAACLRSQ- 2126
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 186 NGLAAGGYDLMGVDLYQOEVRAPRAFLVNLDSBEAEVSRQHGTWLSPAPCNCSEET 245
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2127 --ASSDSISLSSLSIKSGSIGLSP--FHUTPDQBEKPTNSK------GRPLKPGKS 2173
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 246 TLETAKVVMQWSY---QEWNSLLASSRYSBESFTVUFRQPFYETPSLSEDPRQLD 300
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2226 GRTMII-HIPG--VRNSSLSTSPVS-KKCPPLKTPASKSP-SEGQATATSPRGAKPSVKS 2279
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 361 EV---REGAEIRCPDKPS---DTVPTSVRKLKADINVIGALGDSLTLGNGAGSTP 411
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2280 ELSPVAROTSQIGGSKAPSRRSSRDSRSPSRPAQ-QPLSRPQSPGRNISIQRNGTSP 2338
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 412 GNTLVLDLVTQYRGSLSWVGSDENTIGTVTLANIREFENSLKGE-SVGGK--ETSPN AFL 468
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2339 NK-----LSOLPRTSSPSTASTKSGGKMSYTSpgROM 2372
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 469 NQAVGAGRAEDLIPQVQARRLVLUDMKNDTRIHFOED-WKITLFIGNDLCDFCNLDLVHSP 527
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2373 SQ-----QNLTKQ---TGLSNKASSIPRSSEASKGLNQMNNG---ANKKVELSR 2417
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 528 QNFIDNIGKOLDIHAEVRAFYNLVTLVIMRLYQEKKVYCPRMILRSCLPCVULKF 587
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2418 MSSTKSSSESDD--RSERP-----VLRQSTFIKE-----APSPTRL----- 2452
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 588 DDNSTLATLIEFNKKFQBKTQHLESGRYDTRDFTVWQPFENVMPKUTSEGLPNS 647
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2453 -----RKLESASFESLSPSSRASPRTSQAOTPVLSPSLPD-- 2489
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 648 FFAPDCFFHSSSKSHSRAASALWNMM--LEPV---GOKTRHKPFENKINITCPNQVQ 699
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2490 -----MSLSTHSSVQAGGWWRKLPNPLSPTEBYNDRPAKHDIA-RSHSSPSRL-P 2539
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 700 FLRUYKKNMOSHGHFLWPCDRDRASLAHPTSVHALRPADQVVAAGDILTAGNGISKD 759
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2540 INRS-----GIWKREHSKHSSSLPRVSTWRRTGSSSLSASSESEK----AKSE 2586
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 760 DLPUVTTQYRGLEYISAGGDSLENTVTLNLR---EPM RNLTGVAVGIGDANDTNA 813
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2587 DERKWS-----ISGTRQSKENQVSAKGTRWKIKENEFSPTNSTSQVSSCATN--- 2635
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 814 FLNOAVGAPGKAEDIMISQVOTLWQKMD 840
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2636 -----GAESKTLIQMAPAVSKTD 2655

RESULT 12
 US-07-741-940-2
 ; Sequence 2, Application US/07741940
 ; Patent No. 5352775
 ; GENERAL INFORMATION:
 ; APPLICANT: ALBERTSEN, HANS
 ; APPLICANT: ANAND, RAKESH
 ; APPLICANT: CARLSON, MARY
 ; APPLICANT: GRODEN, JOANNA
 ; APPLICANT: HEDGE, PHILIP J.
 ; APPLICANT: JOSLYN, GROFF

93 QDNTTERPOQVCM--GWTMVLSPDR-----YFSPSVPM----- 123
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 1951 QNFAIENTPVCFSHNSLSSLSIDQENNNKNEPIKEPPDSQGERSKPKQASGAPKS 2010
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2011 FHVEDTPVCFSRNSLSSLSIDSEDDL--LQCISAMPKKPSRLKGDNKHSRNM 2067
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 153 ---PVC---HTGKRVIPHDGAEDLWIAQBLV-----RNM 152
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2068 GGILGEDLTLDLKDQIRDSEHGLSPDSENFDWKAIORGANISIVSLHQAAAACLRSQ- 2126
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 186 NGLAAGGYDLMGVDLYQOEVRAPRAFLVNLDSBEAEVSRQHGTWLSPAPCNCSEET 245
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2128 --ASSDSISLSSLSIKSGSIGLSP--FHUTPDQBEKPTNSK-----GRPLKPGKS 2174
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 246 TLETAKVVMQWSY---QEWNSLLASSRYSBESFTVUFRQPFYETPSLSEDPRQLD 300
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2175 TLETAKTLESESKIGKGSKVVYSLITCKVRSNSEISQMKOP-----LOANMPISR 2226
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 301 STTLAWHLWNRMMEPAGEKDERELUSVKHGRPMKCPQSOPSPYFYSYRNSYLTLCQPKDQL 360
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2227 GRTMII-HIPG--VRNSSLSTSPVS-KKCPPLKTPASKSP-SEGQATATSPRGAKPSVKS 2280
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 351 EV---REGAEIRCPDKPS---DTVPTSVRKLKADINVIGALGDSLTLGNGAGSTP 411
 QY ::::::::::::::::::::: ;:::|||: ;:::|||: ;:::|||: ;:::|||:
 Db 2281 ELSPVAROTSQIGGSKAPSRRSSRDSRSPSRPAQ-QPLSRPQSPGRNISIQRNGTSP 2335

QY 412 GNVLDLVLTQYRGLSWSVGGDENIGNVTILANLIREFNPNSLKGK--ETSPNAFL 468
Db 2340 NK-----ISQLPRSSPSTASTKSSGGSKSYTSFORQM 2373
QY 469 NOAVAGGRAEDLPVQARRLVDMKNDRTRHFQD-WKITLITGGNDLCDRNDLYHYP 527
Db 2374 SQ-----ONLTQ---TGLSKNASSIPRSBESASKGLNQMNNG---ANKVVELSR 2418
QY 528 QNPFDNIGKALDLHAEVPRAFNUVLVLETVNRLIYQEKKVYCRMILSLCPVULK 587
Db 2419 MSSTKSSESD---RSERP----VLRQSTPIKE-----ASPTLR----- 2453
QY 588 DDNSTELATLIEFNKKFOEKTHOLIESGRYDFTVVQPFENNDMPKTSEGPDNS 647
Db 2454 -----RKLEASASPESSLSPSRPASTRSQATPVSPSLP----- 2490
QY 648 FFAPDCFHFSKSHSRASALNNM---LEPV----GOKTRHKFENKINTCPNOVQ 699
Db 2491 -----MSLTHSSVQAGGRKLPPNLSPTEYNDGRPAKRHDIA-RSHSESPRL-P 2540
QY 700 FRTYKNSMOGHGTWICRDRAPSALHPTSHALRADIQVAALGDSLTAGNGISKPD 759
Db 2541 INRS-----GTWKREHSKSSPRVSTWRRTGSSSLSASSSEK---AKSE 2587
QY 760 DYPDVTTQYRGISYSAGGDGSLENVTLPNIR----BNFNRNLTYGAVGTDANDTA 813
Db 2588 DEKHVNS-----ISGTQSKENQVSAGTWRKIKENEFSPNTSQTVSAGTN--- 2636
QY 814 PIANQAVPGAKAIDLMSQVOTLQMKMD 840
Db 2637 -----GAESKLILYQMAPAVSKTED 2656

RESULT 13
US-08-289-548A-2
Sequence 2, Application US/08289548A
Patent No. 5688212
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GROFF
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THILVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, LTD
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4398

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/289, 548A
FILING DATE: 12-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kaga, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107-46943
TELECOMMUNICATION INFORMATION:

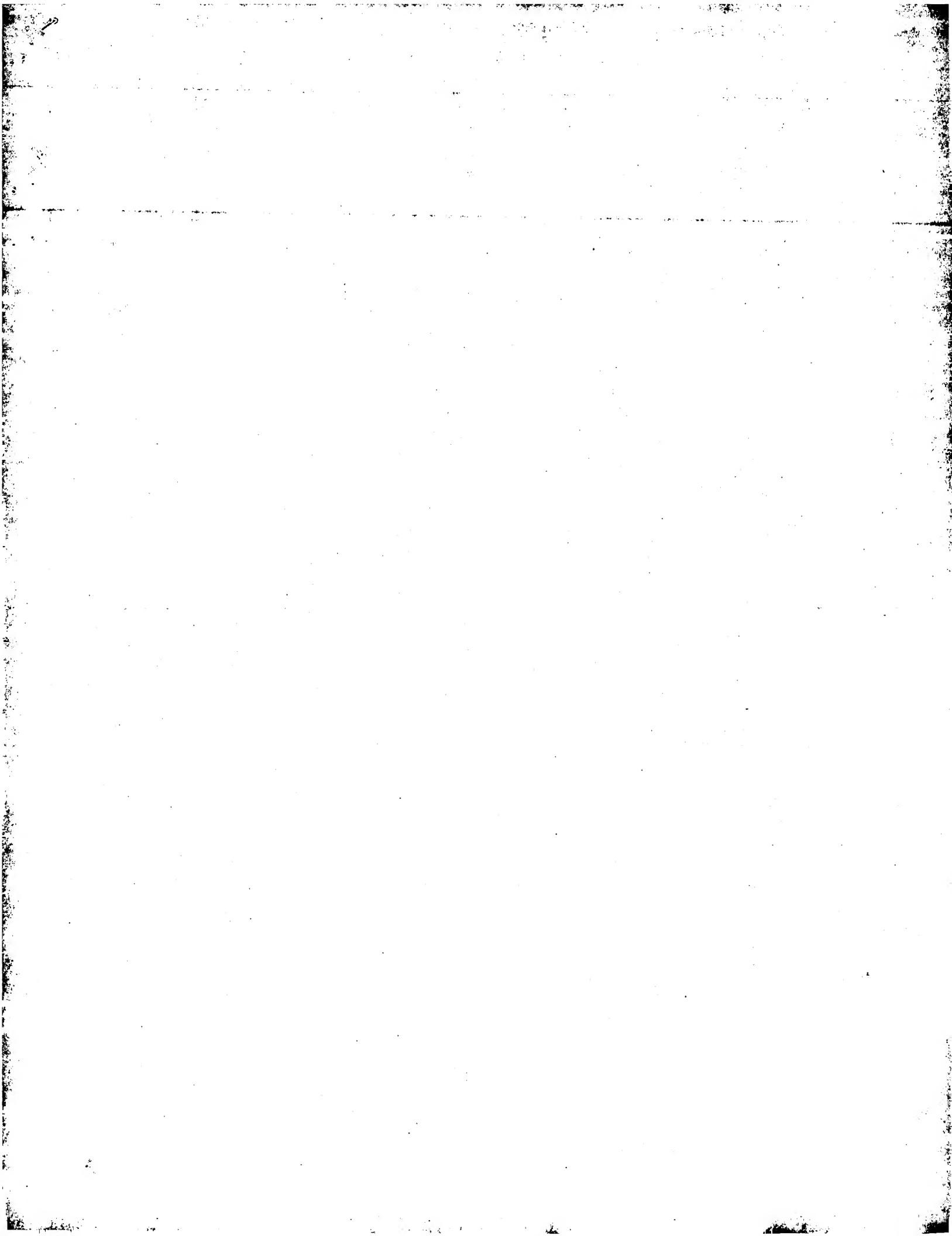
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2843 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-289-548A-2

Query Match Score 1.6%; Score 120.5; DB 1; Length 2843;
Best Local Similarity 19.8%; Pred. No. 0.46;
Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;
Query 34 GOLWPETKNSPPFCPNPKLGVNMPKSVHSLKPSDIKFVAIGNIEIPPPGTCOLEK- 92
Db 1921 GQPKPILQKOSTFP-----QSSKD----- 2540
Query 93 QDWTERPQQVCM---GYMTVLSDIR-----YFSPSPVPM----- 123
Db 1952 QNFAIENTPVCFSHNSSSISDQGNNENEPKITEPPDSQCBPSKPOASGPKS 2011
Query 124 -----PVC---HTGKRVIPHDAEULWIAQOELV-----RNM 152
Db 2012 FIVEDTPVCFSRNSSLSSLSIDEDJL---LOECCISSAMPKKKPSRLKGDNKNEKSPRM 2068
Query 153 -----KENLQLOPRO-----FDWKL---NVFFNSAQ---CYLCPSAQ 185
Db 2069 GGTGLEDLTLDKDIOPDSEHGLSPDSENFDWKA1QBGANSIVSSLHQAACLSRSQ- 2127
Query 186 NGLAAGGVBELMGVLQYLQEVPRAFNUVLVLDSEVAEVRSQYHGTWLSPABPCNCSET 245
Db 2128 -----ASDDSDSTLKSISLQSP---FHLTPOBEKEKFTSNK-----GIRLKQGEKS 2174
Query 246 TRLAKVVMQWSY-----QEWANSLASSRYYQEQESTFVVFOPFFYETTPSILHSEDRLLQD 300
Db 2175 TLETKKIESESEKGKIKGGKVKSLITOKVRNSEISGQMKQ-----LOANMFISR 2226
Query 301 STTLAWLWLNRMMEPAGKEDEPLSVKHGRMKPSOBSPYLFSYRNSNYLTRLQKODKL 360
Db 2227 GRIMI-HIPG--VRNSSSSTSPVS-KKGPKPLKTPASKSP---SEGORTATSPRGAKPSVKS 2280
Query 361-EV---REGAETRCPDKPS---DTVPTSVHRLKPADINVIGALGDSLTAGNGASTP 411
Db 2281 ELSPVARQTSQPGGSSKAPSRSGRSRSTSPTSPRAQ-QPLSRBIOSPNRNSPGRNGISPP 2339
Query 412 GNVLDLVLTQYRGLSWSVGGDENIGNVTILANLIREFNPNSLKGK-SVGTGK--ETSPNAFL 468
Db 2340 NK-----ISQLPRSSPSTASTKSSGGSKSYTSFORQM 2373
Query 469 NOAVAGGRAEDLPVQARRLVDMKNDRTRHFQD-WKITLITGGNDLCDRNDLYHYP 527
Db 2374 SQ-----ONLTQ---TGLSKNASSIPRSBESASKGLNQMNNG---ANKVVELSR 2418
Query 528 QNPFDNIGKALDLHAEVPRAFNUVLVLETVNRLIYQEKKVYCRMILSLCPVULK 587
Db 2419 MSSTKSSESD---RSERP----VLRQSTPIKE-----ASPTLR----- 2453
Query 588 DDNSTELATLIEFNKKFOEKTHOLIESGRYDFTVVQPFENNDMPKTSEGPDNS 647
Db 2454 -----RKLEASASPESSLSPSRPASTRSQATPVSPSLP----- 2490
Query 648 FFAPDCFHFSKSHSRASALNNM---LEPV----GOKTRHKFENKINTCPNOVQ 699
Db 2491 -----MSLTHSSVQAGGRKLPPNLSPTEYNDGRPAKRHDIA-RSHSESPRL-P 2540
Query 700 FRTYKNSMOGHGTWICRDRAPSALHPTSHALRADIQVAALGDSLTAGNGISKPD 759
Db 2541 INRS-----GTWKREHSKSSPRVSTWRRTGSSSLSASSSEK---AKSE 2587
Query 760 DYPDVTTQYRGISYSAGGDGSLENVTLPNIR----BNFNRNLTYGAVGTDANDTA 813
Db 2588 DEKHVNS-----ISGTQSKENQVSAGTWRKIKENEFSPNTSQTVSAGTN--- 2636

ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,655B
FILING DATE: 25-MAY-1995
CLASSIFICATION: 530
PRIORITY APPLICATION NUMBER:
APPLICATION NUMBER: US 08/289,548
FILING DATE: 12-AUG-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/741,940
FILING DATE: 0-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107.43964
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2843 amino acids
TYPE: amino acid
TOPLOGY: linear
MOLECULE TYPE: protein
US-08-452-655B-2

Query Match 1.6%; Score 120.5; DB 1; Length 2843;
Best Local Similarity 19.8%; Pred. No. 0.46; Mismatches 328; Indels 311; Gaps 47;
Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;
QY 3.4 QDWRPETLKNSPFPFCNPNKLGVNMPKSVHSLKPSDKFVAIGNBLIPPDGTGDLBK- 92
OY 1921 GOKPLQKOSTFP-----QSKDI-----PDGAATBKL 1951
Db 124 -----PVC---HTGKRVIPHGDAGELWLIQIQELV-----RNM 152
Qy 93 QDWTRRQQYCM---GVMTVLSDIR-----YFSRSPVPM----- 123
Oy 1952 QNFAENTTPVCFSHNSLSSLDIDEBNNKNEPIKETEPPDSQGPSPKGQASGVAPKS 2011
Db 153 ---KENLODFQ-----FDWKLI---NFFSNASQ---CYLCPSAQ 185
Qy 2069 GGIIGEDLTIDLKD1QRPDSEHGLSDSENFDWKA1QEGANSIVSIIHQAAAACJLRO- 2127
Db 186 NGLAAGGVDELMGVYDYLQEVPRAVNLVLSEVAEVRSQYHGTWLSAPPNCNCBET 245
Qy 2128 ---ASSDSDSILSLSKGISLGSPP---FHLLPDBEKEPTSNK-----GPRILKGCRKS 2174
Db 246 TRLAKVMQNSY-----QEAWSNLLASSRYEQESTFVQPFYETPSLJHSEDPRIQD 300
Qy 2175 TLETKKWLNRMMEPGEKOBPLSVHGRPMKCPSPSBDSPYSPFLSYRNNSYLTLPQKDPKL 360
Db 301 STTLAWHLWNRMMEPGEKOBPLSVHGRPMKCPSPSBDSPYSPFLSYRNNSYLTLPQKDPKL 360
Qy 2227 GRTM-HIPG--VRUSSSSSPVS-KKGPPLKTPASKSP--SEGQMTTSPPRGAKPSVK 2280
Db 361 EV---REGABIRCPDKDPS---DVTUPTSVRILKADINVIGALGDSLTTAGNGGSTP 411
Qy 2281 ELSPVVARQTSQIGGSSKAPSRSGRSRDSTSPPAQP-QPLSPRQISPGRNRNSISPGRNGISIAPP 2339
Db 412 GNVLDVLTQYRGLSNSVGGDNEIGNTTILRENPNSLKG-FWGTGK-BTSPNAFL 468

Search completed: January 6, 2004, 19:37:43
Job time : 44 secs



Copyright (c) 1993 - 2004 CompuGen Ltd.

GanCore version 5.1.6

Om protein - protein search, using sw model

Run on: January 6, 2004, 18:14:14 ; Search time 100 Seconds

(without alignments) 3762.407 Million cell updates/sec

Title: US-10-054-691-2

Perfect score: 7766

Sequence: 1 MGILRGIFFLLELLIGQGT... RCRGGRREDPPMSLRTVAL 1458

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 0

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : SPTRMBL 23::*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_orangelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp Rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID % Query

No. Score Match Length DB ID Description

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|--------|-------|--------|----|---------|---|
| 1 | 5644.5 | 72.4 | 1463 | 11 | OT0320 | 070320 PRELIMINARY; PRT: 1463 AA. |
| 2 | 5229.5 | 68.0 | 1450 | 11 | OT04728 | 070320; Created) |
| 3 | 2270 | 29.2 | 423 | 4 | OBTUP7 | 01-AUG-1998 (TREMBLrel. 07, last sequence update) |
| 4 | 1392 | 17.9 | 270 | 4 | Q96DP9 | 01-MAR-2003 (TREMBLrel. 23, last annotation update) |
| 5 | 1145 | 14.7 | 394 | 11 | Q9D4Y6 | DR Phospholipase B. |
| 6 | 635 | 8.2 | 377 | 5 | Q9H224 | DR |
| 7 | 628.5 | 8.2 | 186 | 11 | Q9K255 | DR |
| 8 | 628.5 | 8.1 | 382 | 5 | Q23119 | DR |
| 9 | 625.5 | 8.1 | 374 | 5 | Q01300 | DR |
| 10 | 616 | 7.9 | 398 | 5 | Q8MxQ8 | DR |
| 11 | 584 | 7.6 | 424 | 5 | Q9W80 | DR |
| 12 | 515 | 6.5 | 447 | 5 | Q9VMM8 | J. Biol. Chem. 273:13407-13414 (1998). |
| 13 | 469.5 | 6.0 | 460 | 5 | Q950L1 | DR |
| 14 | 406.5 | 5.4 | 348 | 5 | Q21799 | DR |
| 15 | 360.5 | 4.6 | 528 | 5 | P98862 | DR |
| 16 | 360.5 | 4.6 | 565 | 5 | Q814J6 | DR |

| | | | | | | |
|----|-------|-----|-------|----|--------|----------------------|
| 17 | 359 | 4.6 | 484 | 5 | Q21798 | 021798 caenorhabdi |
| 18 | 314 | 4.0 | 471 | 5 | Q9n4P5 | Q9n4P5 caenorhabdi |
| 19 | 215 | 2.8 | 433 | 5 | Q965W0 | Q965W0 caenorhabdi |
| 20 | 159.5 | 2.1 | 941 | 16 | Q97C23 | Q97C23 clostridium |
| 21 | 146.5 | 1.9 | 1755 | 16 | Q9PC33 | Q9pc33 xylophaga fab |
| 22 | 145 | 1.9 | 2153 | 12 | Q9YQ05 | Q9YQ05 tula virus. |
| 23 | 141.5 | 1.8 | 799 | 4 | Q8N265 | Q8n265 homo sapien |
| 24 | 141.5 | 1.8 | 873 | 5 | Q01922 | Q01922 caenorhabdi |
| 25 | 139 | 1.8 | 9439 | 16 | Q8CP76 | Q8cp76 staphylococ |
| 26 | 137 | 1.8 | 4306 | 11 | Q9JJ79 | Q9jj79 rattus norv |
| 27 | 136 | 1.8 | 685 | 16 | Q8XST4 | Q8xst4 ralstonia s |
| 28 | 135 | 1.7 | 1399 | 5 | Q9V4R0 | Q9v4r0 drosophila |
| 29 | 134.5 | 1.7 | 16274 | 5 | Q8IR22 | Q8f22 drosophila |
| 30 | 133 | 1.7 | 1708 | 17 | Q26769 | Q26769 methanobact |
| 31 | 132 | 1.7 | 809 | 3 | Q9HEA5 | Q9heas neurospora |
| 32 | 132 | 1.7 | 181 | 16 | Q9PLLB | Q9px31 chlamydia m |
| 33 | 131.5 | 1.7 | 1003 | 10 | Q948S6 | Q948s6 nicotiana t |
| 34 | 131 | 1.7 | 465 | 12 | Q90729 | Q90729 human papil |
| 35 | 130.5 | 1.7 | 891 | 10 | Q9F240 | Q9f240 arabidopsis |
| 36 | 130.5 | 1.7 | 1165 | 5 | Q01923 | Q01923 caenorhabdi |
| 37 | 130.5 | 1.7 | 1545 | 5 | Q9XX31 | Q9xx31 caenorhabdi |
| 38 | 130.5 | 1.7 | 2792 | 5 | Q914R2 | Q9irz plasmoidium |
| 39 | 130.5 | 1.7 | 2922 | 5 | Q8IPJ3 | Q8ipj3 drosophila |
| 40 | 130.5 | 1.7 | 3270 | 4 | Q8WWW3 | Q8ww3 home sapien |
| 41 | 130 | 1.7 | 923 | 17 | Q8U434 | Q8u434 pyrococcus |
| 42 | 129.5 | 1.7 | 1245 | 12 | Q9YTQ9 | Q9ytq9 atelie her |
| 43 | 128.5 | 1.7 | 974 | 10 | Q8905 | Q8s905 arboidopsis |
| 44 | 128.5 | 1.7 | 1003 | 10 | Q9LPQ5 | Q9lpq5 arabidopsis |
| 45 | 128.5 | 1.7 | 2765 | 11 | Q9R1K2 | Q9rik2 rattus norv |

Database : SPTRMBL 23::*

ALIGMENTS

| | | | | | | |
|-----|-------------------|---------------|---------------|---------------|---------------|---------------|
| 1: | sp_archaea:* | SPTRMBL 23::* |
| 2: | sp_bacteria:* | SPTRMBL 23::* |
| 3: | sp_fungi:* | SPTRMBL 23::* |
| 4: | sp_human:* | SPTRMBL 23::* |
| 5: | sp_invertebrate:* | SPTRMBL 23::* |
| 6: | sp_mammal:* | SPTRMBL 23::* |
| 7: | sp_mhc:* | SPTRMBL 23::* |
| 8: | sp_orangelle:* | SPTRMBL 23::* |
| 9: | sp_phage:* | SPTRMBL 23::* |
| 10: | sp_plant:* | SPTRMBL 23::* |
| 11: | sp_Rodent:* | SPTRMBL 23::* |
| 12: | sp_virus:* | SPTRMBL 23::* |
| 13: | sp_vertebrate:* | SPTRMBL 23::* |
| 14: | sp_unclassified:* | SPTRMBL 23::* |
| 15: | sp_rvirus:* | SPTRMBL 23::* |
| 16: | sp_bacteriap:* | SPTRMBL 23::* |
| 17: | sp_archeap:* | SPTRMBL 23::* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

QY 61 SVSILKPSDIFKAIGNLEIIPDPGPGDLEKODWTER-POQVCMGVMTLSIDIRYFP 119
 Db 60 SVHILTPADIKLAIAIGMETPPDGSATNLDSBTERCEPWRGCMGMVNTLSDISHEP 119
 QY 120 SVMPVCHTGKRVPHDGAEIDLWTOAQELVRNKKENLOLDFFQFDWKLNINVFFNSACOLY 179
 Db 120 SVILPTCPWRSAAVRGGVEELTQAELVSIKANPQLDFQODWKLNINVFFNSACOLY 179
 QY 180 CPEAQONGLAAGGYDELMGLYDLYQEVPRAFNLVUDSEVAEVSRQHTGMIS-PAPER 238
 Db 180 CPSAHENGPLMSNNKDLAGILHYLHQEVPRAFNLVUDPEVAMPRHGTWLSRSPEA 239
 QY 239 CNCEBETTRLAKVOMWSYQARNLASSRYSRQEESTVUOPPFYETTPSLHSEPREL 298
 Db 240 CGSGSETSKLDTVUNWQSYOETMDS 297
 QY 299 QDSTTLAHLWNRNMEPAGEBKEPLSVKHGRPMKCPSCQESPYIFSYRNNSYJLQKRD 358
 Db 298 QDTTLAISWNMMKPGQKDPPFSTERPKCPSCQESPYIFSYRNNSYJLQKRD 357
 QY 359 KLEVREGABIRCPCDKDSDTVPSPVSRHKPDINVIGALGPDLTAGNCAGSPGNVUDL 418
 Db 358 QHKREGIEBIRCPDKDSDTVPSPVSRHKPDINVIGALGPDLTAGNCAGSPGNVUDL 417
 QY 419 TQYRGLSNSVGGENIGNYTTLANIREFNPNSLJKGFSVGTGKETSPNAFLN2AVAGRAE 478
 Db 418 TYRGLSNSIGADHNISVTTLNIRLFNPNSLJKGFSVGTGKANSGAFFENAVAGAR 477
 QY 479 DLVQARSLBLVDLMKNDTRTHFEDWKILITLFFGNDLCDFCIDLVHSPONTDNIGAL 538
 Db 478 DLVQARLTIVDLMQHTSINFEDWKITIVFIGNDLCDFCSDPVTSNPSENTDNIGAL 537
 QY 539 DILHAEVRAFYNLVLUFBINIRELYQEKKYVCPRMLRLSICPCVLUFDINSTELATL 598
 Db 538 DILHAEVRAFYNLVLUFBINIRELYQEKKYVCPRMLRLSICPCVLUFDINSTELATL 597
 QY 599 EPNKKPQKETHQLESGRYDTRDFTVUOPFENVDMPKTSRGLPONSPAPDCPFHRSS 658
 Db 598 DINKKQYBERTHOLIESGRYDTRDFTVUOPFEEVKDPIKTEBGLPONSPAPDCPFHRSS 657
 QY 659 KSHSRAASALWNMLEPGQKTRHKFKENKINTCPNOVQPERTYKNSMOHGHTWLPCR 718
 Db 659 KTHARASALWNLQLEPGQKTRHKFKENKINTCPNOVQPERTYKNSMOHGHTWLPCR 717
 QY 719 DRAPSALHPTSAHLRADIQVVALGDSLTAAGTGSKSPDLPDVTOYRQSYNSGGD 778
 Db 718 ERTPSASPSTSAHLRADPVVVALGDSLTAAGTGSKSPDLPDVTOYRQSYNSGGD 777
 QY 779 GSELENVUTPLNIREFENLTAVGTCANDTNAFLQAOVQGAKADLMSQVOTMOK 838
 Db 778 GSLENVUTPLNIREFENLTAVGTCANDTNAFLQAOVQGAKADLMSQVOTMOK 837
 QY 839 KDDDRVNEHDWKVUTVLLGSDLCDXTDSNLYSAANFVDIRNALDVLHREPRVLVN 898
 Db 839 KDPKRINHEDWKRIVTULGTDLQHCTDLYSSANFNLNADLHREPRVLVN 897
 QY 899 LQDFLNPTIMROVFLGNGPDKCPCVQASYLICNCTLTBNSOLARLARFSRYSRSSL 958
 Db 898 LVDFMNPSIMRQVFLGNGPDKCPCVQASYLICNCTLNSYLRENSYLARMDALTRAYQSSREL 957
 QY 959 VSSGRYDQTOEDSFVUQPFQNLQPLVADGPDTSFAPDCIPNQKHSQKLSRALWTN 1018
 Db 958 VSSGRYDQFRESVNLQPFQNLQPLVADGPDTSFAPDCIPNQKHSQKLSRALWTN 1017
 QY 1019 MIEPLGSKETDUDRAEMPTCPTQNEBFLRPRNSNTYTKPALENWGSPLCTEWKA 1078
 Db 1018 MIEPLGSKETDUDRAEMPTCPTQNEBFLRPRNSNTYTKPALENWGSPLCTEWKA 1077
 QY 1079 SNSVPTSYHOLRPAIDIKVVAAGLDSLTAVGARPNNSSDLPSPWGLSWSGGDNLETH 1138
 Db 1078 SNSVPTSYHOLRPAIDIKVVAAGLDSLTAVGARAJSNDSLMSWRLGSWSICGDGALETH 1137
 QY 1139 TTPNPLKKFNPVYLGFSTSTWEGTAGLNVAEGARARMDMPAQAWDIVERMNSPDINL 1198

RESULT 2
 ID 054728 PRELIMINARY; PRT; 1450 AA.
 AC 054728
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Phospholipase B.
 OS Rattus norvegicus (Rat).
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SD; TISSUE=Small intestine;
 RX MEDLINE=98113187; PubMed=9442065;
 RA Takemoto H., Zolotaryov F.N., Ting L., Urbain T., Komatubara T.,
 RT "Identification of functional domains of rat intestinal phospholipase
 B/lipase. Its cDNA cloning, expression, and tissue distribution.";
 RT J. Biol. Chem. 273:2222-2231(1998).
 DR EMBL; D3648; BAB2813.1;
 DR InterPro; IPR001087; Lipase_GDSL.
 DR Pfam; PF00557; Lipase_GDSL; 3.
 DR PROSITE; PS01098; LIPASE_GDSL_SER; 2.
 SQ SEQUENCE 1450 AA; 16108 MW; 455589C8FD91F45 CRC64;
 Query Match 68.0%; Score 5279.5; DB 11; Length 1450;
 Best Local Similarity 70.0%; Pred. No. 0;
 Matches 1012; Conservative 161; Mismatches 263; Indels 9; Gaps 6;
 Matches 1012; Conservative 161; Mismatches 263; Indels 9; Gaps 6;
 QY 5 PGFL--LELLULLLGQCTPQHTSPRKSTLQGLWPETKNSPPCPNPKLGVNMPSKV 62
 Db 5 PGSVLSVLLVLLLGQCTPQHTSPRKSTLQGLWPETKNSPPCPNPKLGVNMPSKV 64
 QY 63 HSLKPSDIFKAIGNLEIIPDPGTC--DLK-QDPTERROQVCMVMTLSIDIRYFP 119
 Db 65 HSLRPSDIFKLVAIGNLEIIPDPGTC--DLK-QDPTERROQVCMVMTLSIDIRYFP 124
 QY 120 SYMPVCHTGKRVPHDGAEIDLWTOAQELVRNKKENLOLDFFQFDWKLNINVFFNSACOLY 178
 Db 125 SYLMPICSPGKTAGHTIADLWTOAQELVRNKKENLOLDFFQFDWKLNINVFFNSACOLY 184
 QY 179 LCPAQONGLAAGGYDELMGLYDLYQEVPRAFNLVUDSEVAEVSRQHTWLSPAPER 238
 Db 185 LGSDQQRHLMKMEMLSGLDLYLREVRPFAVNUVLDSEVLTMAQQHBTGFSAPEI 244
 QY 239 CNCEBETTRLAKVOMWSYQARNLASSRYSRQEESTVUOPPFYETTPSLHEDDPR 298

| | | | | | |
|----|--|------|----|--|------|
| Db | 245 CKCSEETKLSIKAQMOSYQEAWEILLASSKPKHETFAWVQSFSSPV-BLPLRSP | 302 | Oy | 1378 NFTHSRAKLKCOPSPESPYLYTURNSRLPQDSEEAPEBVLYAVPVAGVGLVNGI1GTVV | 1437 |
| Oy | 299 QSTTLAHLWNRUMPERAGEKEDEPLSVKGGRPKMPCPQESPSPLFSYNSNTYTRLOKPD | 358 | Db | 1383 NFTYNRKLKCOPSPENPFLYTVRSQILDKRAKENSNTLYAVPVAVGGLVNGML | 1442 |
| Db | 303 QSTTLAHLWNRUMPERAGEKEDEPLSVKGGRPKMPCPQESPSPLFSYNSNTYTRLOKPD | 362 | Oy | 1438 WCCR 1442 | |
| Oy | 359 KLEVREGAEIRCPDKDSDTIVPTSVHRLKPADINVIGALGDSITAGNQAGSTPGNVLD | 418 | Db | 1443 WRTVR 1447 | |
| Db | 363 KQMKEGTKFTCPDKDSDTIVPTSVHRLKPADINVIGALGDSITAGNQAGSTPGNVLD | 422 | | | |
| Oy | 419 TQYRGLSHVGSDENIGVTTJANILREFNPSSKGFSVGTGKETSPNAFLNGAVAGRAE | 478 | | | |
| Db | 423 TQYRGLSHVGSDETIERTVTTJANILREFNPSSKGFSVGTGKETSPNAFLNGAVAGRAE | 482 | | | |
| Oy | 479 DLVQARQLVLQDMNDTRHFDWDWKITLTFTGNDLCCDFCILVHYSPONTDNICKAL | 538 | Db | 081UP7 PRELIMINARY; PRT; 423 AA. | |
| Db | 483 GLAQAKKLVLVSKMDKDTINFOBDWKITVFGLDLCGSCNLARFSPOTIDNIKAL | 542 | Oy | 081UP7 PRELIMINARY; PRT; 423 AA. | |
| Oy | 539 DLHAEVPRAFNLYTVEIWNRELVOEKOVYCPMLRSJCPVCVLFDDNSTELATLI | 598 | Db | 01-MAR-2003 (TREMBrel. 23, last sequence update) | |
| Db | 543 DLHAEVPRAFNLYVSBIEPLRELNEPEKVSICPRMLSRCPVNLGENSAELQV | 602 | Oy | 01-MAR-2003 (TREMBrel. 23, last annotation update) | |
| Oy | 599 EPNKKFQEKTHQJLESGRYDTEDFTVVWQPFFENVPDKTSBGLPDNSFFADCFHFS | 658 | Db | Similar to phosphatase B. | |
| Db | 603 ENRQYQETGKLEISGVYDFTVVLQPMENVNMPRTLEGULPSSFFADCFHFS | 662 | Oy | Homo Sapiens (Human). | |
| Oy | 659 KHSRASAALWNNLRLPVGKUTRKHENKINTCPNOVOPRTYKNSMOHGWTPLCR | 718 | Db | Bukaryota; Metacozia; Craniata; Vertebrata; Euteleostomi; | |
| Db | 663 KTHARSAALWNNLRLPVGKUTRKHENKINTCPNOVOPRTYKNSMOHGWTPLCR | 722 | Oy | Mammalia; Buthezia; Primates; Catarrhini; Homindae; Homo. | |
| Oy | 719 DRAPSALHPTSHALRADIQVAALGDSLTINGGIGSKPDDLPDVITYQRLSYSGGD | 778 | Db | NCBI_TAXID=9606; | |
| Db | 723 ERKPSASRPTSVHLTRADIQVAALGDSLTINGGIGSKPDDLPDVITYQRLSYSGGD | 782 | Oy | [1] | |
| Oy | 779 GSLENVTLPNIREFERNLT3YAVGTDANDTNAFLINQAVGAKADEDMSQVLMORM | 838 | Db | SEQUENCE FROM N.A. | |
| Db | 783 KULEANVTPNIREFERNLT3YAVGTDANDTNAFLINQAVGAKADEDMSQVLMORM | 842 | Oy | RC | |
| Oy | 839 LUDFLNPTIMRQFLGDPCKPVQQAStLNCULTURENSQSLARLEAFSRAYRSSREL | 958 | Db | RA | |
| Db | 903 LUDFMNPSIIRQYFLKQDPCPNOTSQYCNLTCVLPBDHSHLARFAFTKSQSMQL | 952 | Oy | Straubberg R.; | |
| Db | 843 KUDTRVNFHQDWKIVTMIGASDLCDFCKDSKRYSAANFSDHLMRNLDLHKEVPLRVN | 902 | Db | RU | |
| Oy | 899 LUDFLNPTIMRQFLGDPCKPVQQAStLNCULTURENSQSLARLEAFSRAYRSSREL | 958 | Oy | DR | |
| Db | 903 LUDFMNPSIIRQYFLKQDPCPNOTSQYCNLTCVLPBDHSHLARFAFTKSQSMQL | 952 | Db | EMBL; BC042674; AAH2674_1; - | |
| Oy | 959 VSEGRYDQTBDSVVLQFFQNLQLPVLADGPDTSFAPDCITHPNOKFHSQALARLWTN | 1018 | Oy | A20A193491BA780C CRC64; | |
| Db | 963 VSEGRYDQTBDSVVLQFFQNLQLPVLADGPDTSFAPDCITHPNOKFHSQALARLWTN | 1022 | Db | SQ | |
| Oy | 1019 MIEPLGSKTETIPLRAMEPNTCPTONEPFLRPRSNNTYTPKPAIENWSGDFLCTENKASNSVPTSYHOLRPADIK | 1078 | Oy | Query Match Best Local Similarity Score 2270; DB 4; Length 423; | |
| Db | 1023 MIEPLGSKTETIPLRAMEPNTCPTONEPFLRPRSNNTYTPKPAIENWSGDFLCTENKASNSVPTSYHOLRPADIK | 1082 | Oy | Matches 422; Conservative 1; Mismatches 0; Gaps 0; | |
| Oy | 1079 SNSVPTSTHQLRADIKOVALGDSLTTAVGARNPNNSDLPFWWRGLSWSIGDGDLNETHTPLNILKFKNPFLGF | 1138 | Db | 1 MITCPIONEPLRPRSNNTYTPKPAIENWSGDFLCTENKASNSVPTSYHOLRPADIK | |
| Db | 1083 SKVPTSTHQLRADIKOVALGDSLTTAVGARNPNNSDLPFWWRGLSWSIGDGDLNETHTPLNILKFKNPFLGF | 1142 | Oy | 1156 STSWEGSLTTAVGPBNNSDLPTSWRGLSWSIGDGDLNETHTPLNILKFKNPFLGF | 1155 |
| Oy | 1139 TLPNLLKCFNPPULGSTSTWGTAGINVAAGARADMPAQAWDIVERMINSPTNLE | 1198 | Db | 1211 STSTWEGTAGINVAAGARADMPAQAWDIVERMINSPTNLE | 180 |
| Db | 1143 TLPNLLKCFNPPULGSTSTWGTAGINVAAGARADMPAQAWDIVERMINSPTNLE | 1202 | Oy | 61 WVALGDSLTTAVGPBNNSDLPTSWRGLSWSIGDGDLNETHTPLNILKFKNPFLGF | 120 |
| Oy | 1199 KWKWLVLFIGNDLCHYCENPAHALTEYVHIOQDILSERELPRAFYUNVEMELASLYQCGGKCAMLAQQN | 1258 | Db | 1156 STSWEGTAGINVAAGARADMPAQAWDIVERMINSPTNLE | 1215 |
| Db | 1203 EPMKLITLIGNDLCLYCEDPENYSTREYVYKIOHALDIFYELPRLPVFINVEMBLG | 1262 | Oy | 1216 YCENPEAHALATEYVHIOQDILSERELPRAFYUNVEMELASLYQCGGKCAMLAQQN | 1275 |
| Oy | 1259 LVOGQOGGKCAMLAAQNCTCIRHSQSLEKDLKQWNLQHGISSFSYMQYQTQDF | 1317 | Db | 181 YCENPEAHALATEYVHIOQDILSERELPRAFYUNVEMELASLYQCGGKCAMLAQQN | 240 |
| Db | 1263 LWDQGKGCAMPALAVOKNCSCLRSONLAMWBLKQYNGLQALSLSYWHRYMDF | 1322 | Oy | 1276 CTCLRSQSSLEKQELKQYNNLQHGSRSKSYWQHTOREPAFWVQPFONTLPLNR | 1335 |
| Oy | 1318 AWVQPFRTQNTLPLNERGDTDTFFSDFCFSRDRGSHAEMAIALWNNMLPVGKRTSN | 1377 | Db | 241 CTCLRSQSSLEKQELKQYNNLQHGSRSKSYWQHTOREPAFWVQPFONTLPLNR | 300 |
| Db | 1323 AVVQPFRTQNTLPLNERGDTDTFFSDFCFSRDRGSHAEMAIALWNNMLPVGKRTSN | 1382 | Oy | 1336 GTDTLPFSEDCPHFSRDRGSHAEMAIALWNNMLPVGKRTSN | 1395 |
| Oy | | | Db | 301 GTDTLPFSEDCPHFSRDRGSHAEMAIALWNNMLPVGKRTSN | 360 |
| Db | | | Oy | 1396 LYTLRNSRLLPQDAESEAPEVYAWPVAGVGLVNGITGVWRCRRGGRBDDPPMSLT | 1455 |
| Oy | | | Db | 361 LYTLRNSRLLPQDAESEAPEVYAWPVAGVGLVNGITGVWRCRRGGRBDDPPMSLT | 420 |
| Oy | | | Oy | 1456 VAL 1458 | |
| Db | | | Db | 421 VAL 423 | |
| | RESULT 4 | | | | |
| | 096DP9 | | | | |
| | ID | | | | |
| | 096DP9 | | | | |
| | PRELIMINARY; | | | | |
| | PRT; | | | | |
| | 270 AA. | | | | |
| | AC | | | | |
| | 096DP9; | | | | |
| | DT | | | | |
| | 01-DEC-2001 (T-EMBLrel. 19, Last sequence update) | | | | |

DT 01-OCT-2002 (TREMBrel, 22, Last annotation update)

DE Hypothetical protein FLJ30866.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxId:9606; [1]

RN SEQUENCE FROM N.A.

RP TISSUE:Brain;

RA Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Ishii S., RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Nishikawa T., RA Yamamoto J., Isono Y., Matsuo K., Nakamura Y., Sekine M., Kimura K., Yamashita H., Matsuo K., Nakamura Y., Saito K., Nishikawa T., RA Kikuchi H., Kanda K., Waga-Suzuki M., Murakawa K., Kanehori K., RA Sugano S., Nagahashi K., Maeho Y., Nagai K., Iegai T., RA "NEDO Human cDNA sequencing project";

RT Submitted (Oct-2001) to the EMBL/GenBank/PDBJ databases.

DR EMBL; AK055428; BAB7020.1; DR InterPro; IPR001087; Lipase_GDSL.

DR Pfam; PF00637; Lipase_GDSL; 1.

KW Hypothetical protein

SQ SEQUENCE 270 AA; 30869 MW; AC53CB75FD09523 CRC64;

Query Match Best Local Similarity 17.9%; Score 1392; DB 4; Length 270; Matches 259; Conservative 0; Pred. No. 2.2e-95; Indels 0; Gaps 0;

QY 491 MKDTRIRHQBDWKITIFIGNDLICDFNDLPHYSPOQNFTDNGKALDILAEVPRAFV 550

Db 1 MKDTRIRHQBDWKITIFIGNDLICDFNDLPHYSPOQNFTDNGKALDILAEVPRAFV 60

QY 551 NLTVLVEIINRLRELYQEKVYCPMLRSCLPCVLUKFDDNSTELATLIEFNKKFQKTHQ 610

Db 61 NLTVLVEIINRLRELYQEKVYCPMLRSCLPCVLUKFDDNSTELATLIEFNKKFQKTHQ 120

QY 611 LIESGRYDPREDFWVQPFENVDMPTSEGPDNSPAPDCFHSSKSRRASALWN 670

Db 121 LIESGRYDPREDFWVQPFENVDMPTSEGPDNSPAPDCFHSSKSRRASALWN 180

QY 671 NMIEPVGQKTRIKPFENKINTCENQVOPFLRTYKNSQMGHGTWLPCRDRAASALHPTSV 730

Db 181 NMIEPVGQKTRIKPFENKINTCENQVOPFLRTYKNSQMGHGTWLPCRDRAASALHPTSV 240

QY 731 HALRPADIOVVAALGDSL 749

Db 241 HALRPADIOVVAALGDSL 259

RESULT 5

O9D4Y6 PRELIMINARY; PRT; 394 AA.

AC Q9D4Y6; DT 01-JUN-2001 (TREMBrel, 17, Created)

DT 01-JUN-2001 (TREMBrel, 17, Last sequence update)

DE 4330532a06Rik protein.

GN 4330532a06Rik

OS Mus musculus (Mouse)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. OC NCBI_TaxId=1090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=G57BL/6J; TISSUE=Testis;

RC MEDLINE=2105660; PubMed=1217851;

RA Kawai J., Shinagawa A., Shihata K., Yoshino M., Itoh M., Ishii Y., RA Azawa K., Hara A., Nishii K., Kiyoshi H., Kondo S., Yamana I., RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., RA Kadota H., Matsuda H.A., Ashburner M., Batyalov S., Casavant T., RA Fleischmann R., Gaasterland C., Gissi C., King B., Kochiba H., RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., RA Blake J., Boffelli D., Bojunga J., Carninci P., de Bonaldo M.F., RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., RA Gustincich S., Hill D., Hoffmann M., Hume D.A., Kamiya M., Lee N.H., RA Lyons P., Marchionni L., Mashima J., Mazzairelli J., Monbaerts P., RA Nordone P., Ringwald M., Rodriguez T., Sakamoto N., RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibaoka Y., Storch K.-F., RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittraker C., Wilming L., RA Wynnshaw-Borisi A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S., RA Hayashizaki Y.;

RT "functional annotation of a full-length mouse cDNA collection." Nature 409:685-690 (2001).

DR EMBL; AK015993; BAB30072.1; DR MGI; MGI:1922406; 4930539a06Rik.

SQ SEQUENCE 394 AA; 44290 MW; E6DB3EC17E8F0888 CRC64;

Query Match Best Local Similarity 14.7%; Score 1145; DB 11; Length 394; Matches 234; Conservative 46; Mismatches 97; Indels 14; Gaps 7;

QY 1 MGURPGI--FLERLLIIGGGRPOHNSPPRASLIEQG---LMPETLKNSPPPCNPKLG 54

Do 1 MELYPGI--FLERLLIIGGGRPOHNSPPRASLIEQG---LMPETLKNSPPPCNPKLG 58

QY 55 VAMPSKSVHSKLPSDKRVAATGNLEIPDPGTG--DLIEK-ODWTERPQQVCMGVMTLS 111

Do 59 LSTLSESHVSILPSKPSDIKVLAIGNPEPLAPSGTINMEKOSIQNPQDVGMIVLS 118

QY 112 DIRYFSRSVPVPGVCHGKRVPIPHGDAEDLWQAOELVRVNRMKENLQDFDWKLINVFF 171

Do 119 DILRHFSPSVLMPTCSPEGKTAHTWADLWQAOELVRVJUDNPQDFDWKLITVFF 178

QY 172 SWASQCYTCPSAQONGLAAAGGUDELKGWLVQEQYPRAFNLVDLSEAVASVROQHGTW 231

Do 179 SWTSQCHCPSAQOKSHMRHEMLNGVLDVYHLHVEYPRAFNLVDLSEBVLMQDQETG 238

QY 232 LSPAPEPCNCSEBTTRILAKVQWQS1QEWANSLASSRYSQBSFTVFQDPFYETPSL 291

Do 239 FSAPEPCKT-ETTLISAKVQWQS1QEWADLASSKFNKHETPAVVFQDPFYETPSL 297

QY 292 HSEDPRLDGSTTLAWHWRNRMMEPAGEKEDEPLSVKGRPMKCCPSQESPYLISYRNNTLT 351

Do 298 KRSSP--QDPTTALRIRWNSMHPVGDGLNTABERKTMKPCPSEESPYLYTQNSYQA 355

QY 352 RIQKQPQKLEVEGAERCPDPDSPVPTS 382

Do 356 RUKPPIRQLQMGKSBETCPDKPNSIPTT 386

RESULT 6

O9N2Z4 PRELIMINARY; PRT; 377 AA.

AC O9N2Z4; DT 01-OCT-2000 (TREMBrel, 15, Created)

DT 01-OCT-2000 (TREMBrel, 15, Last sequence update)

DE 01-MAR-2003 (TREMBrel, 23, Last annotation update)

DE Hypothetical 41.4 kDa protein.

GN 165B4B.1.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC NCBI_TaxId=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RC MEDLINE=99069613; PubMed=9851916;

RA None;

RT "Genome sequence of the nematode *C. elegans*: a platform for investigating biology. The *C. elegans* Sequencing Consortium.", Science 282:2012-2018(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;
 RA Maupin R.; Bauer C.; elegans cosmid Y65B4BR."
 RT "the sequence of C. elegans cosmid Y65B4BR."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDJB databases.
 RN [3]
 RP SQUENCE FROM N.A.
 RC Waterston R.;
 RT "Direct Submission."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DDJB databases.
 DR WormPep; YSB4BR.1; CB2857.1; -.
 DR InterPro; IPR001087; Lipase_GDSL.
 DR Pfam; PF00557; Lipase_GDSL; 1.
 KW Hypothetical protein
 SQ SEQUENCE 377 AA; 41400 MW; 2A93D3842C80E6A5 CRC64;
 Query Match Best Local Similarity 8.2%; Score 636; DB 5; Length 377;
 Matches 132; Conservative 71; Mismatches 122; Indels 16; Gaps 8;
 DR 376 SPTVTPSWSRHLKPADINVIGAGLGSITAGNGAGSTPGVNLVQDLYTQYGLSNVGDDENIG 435
 Db 377 SKKVVPSAHSVRFADIKVIGALGSDLSITANGAGAQGDPPLAVLQYQGLAFQGGHSLID 95
 OY 436 TTWTTLANILRENPNSLKGSFSGTGKETSPNAF---LNQAVAGGRAEIDLPAVRLLYDLM 491
 Db 97 EHVTVAWNLKKFSPNLMGYSTGIG-- SANVWPSKLNQAVGAEADIDTQARLVQI 153
 OY 492 KDTTRIHOEDKMKITLFLIGNDLDFCNDLPH--YSPQNFTDNGKALDIHAEVPRAP 549
 Db 154 QSHKEIDVKTDWKLINVFAGNDCAVNDRENGPHSKAIWKVNQAVGAEADIDTQARLVQI 213
 OY 550 VLVTVLIVNRELVQEKKVCPRMILRSCLCPVLFEDNTTELATLIEFNKKFOBTH 609
 Db 214 VSITGMFDMMAMRQI-DKDKYFCGDLAYFE-PCBKNDPFDNIDASCHL---YMDAQ 268
 OY 610 QLIESGRVDTREDFTVWVQPFENNDMRTS_EGLPDNSFFADCFHRSSKSNSRAASAL 668
 Db 269 EGDGIFDSTDFTFVWQPFENGTVPPLKPDGEVULDWFPDCFHRSSKSNSRAASAL 328
 OY 669 WNNMLERPGQKTRHKFEN-KINITCPNQVQPLRTYKNS 708
 Db 329 WNNIVQPVGSKHQVNLSDPTIPLNCPDKCPFPRTKNSV 369
 RESULT 7
 O8K255 PRELIMINARY; PRT; 186 AA.
 AC 08K255;
 DT 01-OCT-2002 (TREMBrel. 22, Created)
 DT 01-OCT-2002 (TREMBrel. 22, Last sequence update)
 DT 01-OCT-2002 (TREMBrel. 22, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Bokaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxId:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strauberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DDJB databases.
 DR EMBL; BC031606; AAH31606.1; -.
 KW Hypothetical protein
 SQ SEQUENCE 186 AA; 21471 MW; 56153A626A22315 CRC64;
 Query Match Best Local Similarity 8.2%; Score 635; DB 11; Length 186;
 Matches 116; Conservative 18; Mismatches 26; Indels 0; Gaps 0;
 OY 1289 QELKQKVNVLQHGQSISSEFSYWHQVTOREDFAVWVQPFENFTPLNERGDTDTFFSDCF 1348
 Db 2 QELKQKVNVLQHGQSISSEFSYWHQVTOREDFAVWVQPFENFTPLNERGDTDTFFSDCF 61
 OY 1349 HPSDRGHAEITALWNMLPEPGKTTSNNPFTHSRAKLKCSPSPSYLTYURNRLPDQ 1408
 DR 023119 PRELIMINARY; PRT; 382 AA.
 AC 023119;
 DT 01-NOV-1996 (TREMBrel. 01, Created)
 DT 01-DEC-2001 (TREMBrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBrel. 23, Last annotation update)
 DR W02B12.1; protein.
 GN W02B12.1.
 OS Caenorhabditis elegans.
 OC Caeorhabditis; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peleriniae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Swinburne J.; Ainscough R.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DDJB databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 DR MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C.elegans*: A platform for
 Science 283:2012-2018(1998)."
 RL
 DR Z66521; CAA91393.2; -.
 DR WormPep; YMDAQ 376.
 DR InterPro; IPR001087; Lipase_GDSL.
 SQ SEQUENCE 382 AA; 41772 MW; 6430145BB8525CA3 CRC64;
 Query Match Best Local Similarity 8.1%; Score 628.5; DB 5; Length 382;
 Matches 139; Conservative 57; Mismatches 119; Indels 37; Gaps 8;
 DR 376 SPTVTPSWSRHLKPADINVIGAGLGSITAGNGAGSTPGVNLVQDLYTQYGLSNVGDDENIG 435
 Db 377 SKKVVPSAHSVRFADIKVIGALGSDLSITANGAGAQGDPPLAVLQYQGLAFQGGHSLID 94
 OY 436 TTWTTLANILRENPNSLKGSFSGTGKETSPNAF---LNQAVAGGRAEIDLPAVRLLYDLM 494
 Db 95 EHVTVAWNLKKFSPNLMGYSTGIG-- SANVWPSKLNQAVGAEADIDTQARLVQI 154
 OY 495 TRIFQBDMKITLFLIGNDLDFC---NDLVHVSQNTDNGKALDIHAEVPRAP 549
 Db 155 SELVKWDKMKITLFLIGNDLDFC---NDLVHVSQNTDNGKALDIHAEVPRAP 214
 OY 550 VLVTVLIVNRELVQEKKVCPRMILRSCLCPVLFEDNTTELATLIEFNKKFOBTH 609
 Db 215 VSITGMFDMMAMRQI-DKDKYFCGDLAYFE-PCBKNDPFDNIDASCHL---ESNKQFTDDI 257
 OY 610 Q-----LISGSGYPTREDFTVWVQPFENY-DMPKTSSEGUDNSFFAPCPFH 656
 DR 258 QCGCFGYMAEKDQIQTGFQDNKDDTFVWQPFENGTVPPLPYASPGVDMTFAPCPFH 317
 OY 657 SSKSHSRAASALWNNMLPEPGKTRHKFEN-KINITCPNQVQPLRTYKNS 707
 DR 318 SAYGHAGHMLWNTVQPVKPKSVNLSPSPVGHCSPNCPEPTKNS 369
 DR 318 SAYGHAGHMLWNTVQPVKPKSVNLSPSPVGHCSPNCPEPTKNS 369
 RESULT 9
 OY 001300 PRELIMINARY; PRT; 374 AA.
 AC 001300;
 DT 01-JUL-1997 (TREMBrel. 04, Created)

DT DT 01-OCT-2002 (TREMBIrel. 22; Last sequence update)
 DE 01-MAR-2003 (TREMBIrel. 23; Last annotation update)
 GN F09C8_1 protein.
 OS C. elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderrine; Caenorhabditis.
 OS NEBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McMurray A.A.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99059613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C.elegans*: A platform for investigating biology";
 RT investigating biology.";
 Science 282:2012-2018(1998).
 DR WormPep; F09C8_1; CE31170.
 DR InterPro; IPR001087; Lipase_GDSL.
 DR Pfam; PF00657; Lipase_GDSL_1.
 DR PROSITE; PS00421; Lipase_GDSL_1.
 DR SEQUENCE 374 AA; 41921 MW; 533E0DB7EB3A41EC CRC64;
 SQ
 Query Match 8.1%; Score 625.5; DB 5; Length 374;
 Best Local Similarity 39.8%; Pred. No. 5.3e-38;
 Matches 139; Conservative 60; Mismatches 111; Indels 39; Gaps 9;
 QY 376 SDPTWPTSVHLKPADINVIGALGDSLTAGNGAGSSTPGNVLVDLTYQYRGLSWV 435
 Db 37 SKKVPTVNVNSVRPADIKLALGDSLTAAANGAGAE--DPVAVVQYGLAFOAGSDKTE 94
 QY 436 TVTTLANIREFNPSSLLKGISVGTVGETSPN---AFLNQAVAGGRADLIVQARSLDLM 491
 Db 95 EHVTIPNLKQKYNPDVFGTSNGIG--SPWVETARLNVAMPGANAKDQGQANQIVQOLL 151
 QY 492 KNDTR-INFOEDWKITLFIGGNDICDFCNLDVHSQSPONFTDNICKALDILHAEVYPRAFV 550
 Db 152 QOHTBTVVNMKEDWKLUINIFGGNDICGYCRPKVEDSPSYNCQADTKAOVITYDNYNPRVT 211
 QY 551 NLVTVLEIWLRELYQEKKYCPRLMLRSICPCVWKFDDNSTELATLIEENKKQB--- 606
 Db 212 SLTGMLHLMQLRQT-DTGHWFCQR-LHHBGCQ-----DSNKFDADIR 254
 QY 607 -----KTHOLIESGRYDTRFVVQOPFFENTDMPKTSEGILPDNSFFAPDCHFSSK 659
 Db 255 QACWDYKIEBKQIETDGTBKNDFFYVQPMQDFDTLIPPMENGKPTQKEFAPDEHFSQW 314
 QY 660 SHSRASALWNMMLPEVGOKTRHKFENKI-NITCPNQVOPPLRTYNS 707
 Db 315 GHALVSTYLWNNTLQPGVSKSTVSNMSVPLTACDAAACPFPRTPKNS 363
 RESULT 10
 QMVXQB PRELIMINARY; PRT; 398 AA.
 ID QBVNXQB
 AC QBVNXQB;
 DT 01-OCT-2002 (TREMBIrel. 22; Created)
 DT 01-OCT-2002 (TREMBIrel. 22; Last sequence update)
 DT 01-MAR-2003 (TREMBIrel. 23; Last annotation update)
 DE Hypothetical protein Y73B6BL_37.
 OS C. elegans.
 OC C. elegans.
 OC Rhabditida; Peloderrine; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol_N2;
 RC MEDLINE=019606; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amara P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Waterston R.;
 RA Waterston R.;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=019606; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amara P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Waterston R.;
 RN
 RP SEQUENCE FROM N.A.
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol_N2;
 RA Graves T.;
 RT "The sequence of *C. elegans* cosmid Y73B6BL.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC004197; AAWM4393_1; -.
 DR WormPep; Y73B6BL_37; CE29931.
 DR InterPro; IPR001087; Lipase_GDSL.
 KW Hypothetical protein.
 SQ SEQUENCE 398 AA; 44003 MW; E6027C51E47E2B76 CRC64;
 SQ
 Query Match 7.9%; Score 616; DB 5; Length 398;
 Best Local Similarity 36.0%; Pred. No. 3e-37;
 Matches 133; Conservative 74; Mismatches 120; Indels 42; Gaps 1;
 QY 374 DP----SDPTWPTSVHLKPADINVIGALGDSLTAGNGAGSSTPGNVLVDLTYQYRGLSWV 422
 Db 30 DPSLYKKSQOPTSAHSVRPADIKLALGDSLTAAANGAGAE--DPVAVVQYGLAFOAGSDKTE 89
 QY 429 CGDENITGVTT-----LANTLREFNPSSLLKGISVGTVGETSPN---AFLNQAVAGGRADLIVQARSLDLM 460
 Db 90 GGDSSEDEHVYVASENFTLNEQTSFAIMQIIMQIALLADWMKMKSPNIMQYSTGIG--SANW 149
 QY 468 -----LNOQAVAGGRADLPVYORRLYDLMKNDUTRIHOEDWKIITLFIGGNDICDFCNLDLV 522
 Db 147 EVSKUNQAVPQAQADITQKALVYOMHRIDIDKNDWKLUINVFIGANDMCYCEKE 200
 QY 524 H-YSFQNFTDINGKALDIHLAEVRAFPVNUVTLVILNLRLYQEKKVVCPRMIRSCL 58
 Db 207 NGIHSKANWKQKVNTA1QIKDHLRITVSMGMDRMHQI-BHDKYFCDGLHVE-C 266
 QY 582 PCVULKFDNSTELATLIEFNKKFQEKTHOLIESGRYDTRFVVQOPFFENVMRKT- 64
 Db 265 PCEKKSSFTNDVSQACHL--YMDAQQELOQSGFEDTSDFDTFLQPFENGITPLK 322
 QY 641 EGLPDNSFRAPPDCFHRSKSKSRASALWNMMLPEVGOKTRHKFENKI-NITCPNQVOPPLRTYNS 699
 Db 322 DGEVNLDWFABDCFHPSRLGHANVACHLWNNTVQPGVSKSTVSNMSVPLTACDAAACPFPRTPKNS 388
 QY 700 PLRTYNSW 708
 Db 382 FIRTPKNSV 390
 RESULT 11
 Q9VNR0 PRELIMINARY; PRT; 424 AA.
 ID Q9VNR0
 AC Q9VNR0;
 DT 01-MAY-2000 (TREMBIrel. 13; Created)
 DT 01-MAY-2000 (TREMBIrel. 13; Last sequence update)
 DT 01-MAR-2002 (TREMBIrel. 20; Last annotation update)
 DE CG7365 protein (GH22081p);
 GN CG7365.
 OS Drosophila melanogaster (Fruit fly).
 OC Drosophila; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Aphidoidea; Drosophilidae; Drosophila;
 OC NCBI_TaxID=7221;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=019606; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amara P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Waterston R.;
 RN
 RP SEQUENCE FROM N.A.
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol_N2;
 RA Graves T.;
 RT "The sequence of the nematode *C. elegans*: a platform for investigating biology. The *C. elegans* Sequencing Consortium.",
 RT Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol_N2;
 RA Graves T.;
 RT "The sequence of *C. elegans* cosmid Y73B6BL.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC004197; AAWM4393_1; -.
 DR WormPep; Y73B6BL_37; CE29931.
 DR InterPro; IPR001087; Lipase_GDSL.
 KW Hypothetical protein.
 SQ SEQUENCE 398 AA; 44003 MW; E6027C51E47E2B76 CRC64;

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champs M., Pfeiffer B.D., RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G., RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D., RA Balow R.M., Bauu A., Baxendale J., Bayraktaroglu L., Beasley E.M., RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., RA Borovka D., Botchan M.R., Bouck J., Brokstein P., Brottier P., RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., RA Dobson K., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., RA Durbin R.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W., RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasper K., RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., RA Jalali M., Kalush F., Karpen G.H., Murphy L., Muzny D.M., Nelson D.L., RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., RA Lasko P., Lei Y., Levitsky A.Y., Li J., Li Z., Liang Y., Lin X., RA Liu X., Mattei B., McIntosh T.C., Mcleod T.C., Mcleod M.P., McPherson D., RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J., RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., RA Yeh J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., Venter J.C., RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacieb J.M., RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B., RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J., RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., RA Yeh J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., RA "The genome sequence of *Drosophila melanogaster*.", RA Science 287:2185-2195(2000). RA 12

RA [12]

RA SEQUENCE FROM N.A.

RP STRAIN-BERKELEY;

RC MBDBLINE=20196006; PubMed=10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D., RA Amatanteidae P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champs M., Pfeiffer B.D., RA Miklos G.L.G., Miklos G.L.G., RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G., RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D., RA Balow R.M., Bauu A., Baxendale J., Bayraktaroglu L., Beasley E.M., RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., RA Borovka D., Botchan M.R., Bouck J., Brokstein P., Brottier P., RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., RA Dobson K., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., RA Durbin R.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W., RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasper K., RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., RA Lasko P., Lei Y., Levitsky A.Y., Li J., Li Z., Liang Y., Lin X., RA Liu X., Mattei B., McIntosh T.C., Mcleod T.C., Mcleod M.P., McPherson D., RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J., RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., RA Yeh J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; RA Science 287:2185-2195(2000). RA 12

RA Query Match 7.5%; Score 584; DB 5; Length 424;

RA Best Local Similarity 38.5%; Pred. No. 8.2e-35; Matches 137; Conservative 57; Mismatches 116; Indels 46; Gaps 12;

OY 369 RCPDKDPSDTIVPSVHRLKPADINVAGLGSITAGNAGSGSPGNGVNDLVTQYRGLWSV 428

Db 88 REPR-----PTSVHRLKGPGDIDVIGMAGDQLTAGNGIFAT-----NLHVTVNRGVMSI 139

OY 429 GEDENIGVTTANILREFNPNSIKPGFSVGTKGETSPNAFLNQAVAGGRAEDTIVQARLLV 488

Db 140 GGDYWRKVLTLAPNLKEFNPVLYGAIKDGISTDRSRFDVLAAMSRSRDWPHMAKV 199

OY 489 DLMKNDTRHFEDWKLTFLFGNDLCCDFCIDLVHY-SPOFTD---NICKALDILHA 543

Db 200 RMVQDERPVNMMSDWKLVTLFIGN--DFCTDCYYPEPEKTVWDHBRNMKTYVRLD 256

OY 544 EPRAFYVNLVTEIWLREL-----YQEKKYCPCRMILSRCPVUJKFDNNSTE 593

Db 257 NYPRMLNIVPA--PNRFLNLTGPPICHS-----TURFCPCLM--GKKGQ 302

QY 594 LATLIERKKRQEKTQHOLIESGRYDREDFTVQPFENNQMPKUSQEGLDPNSFRDC 653

DR 303 LDYLEGIMKRWIAKDFBNEFBNT-ETFTINOPPSQFOPPRTRGOTRTPSSEDC 361

OY 654 FRISSKSISRAASALWNNMLEPVGQKT---TRUKPENKINTCQNQVOPPLRTYKNS 707

DR 362 FRLSQRGHAAANSIWNMIEPGHKSGPATLFE--TPRCPSEMRFPLTRENS 414

QY QVMMB ID O9VMMB PRELIMINARY; PRT; 447 AA.

DR QVMMB; (TREMBrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBrel. 13, Last annotation update)

DT 01-JUN-2001 (TREMBrel. 17, Last annotation update)

DE CG11029 protein.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriidae; Drosophilidae; Drosophila.

OC NCBI_TaxID=7227;

RN [11]

SEQUENCE FROM N.A.

RP STRAIN-BERKELEY;

RC MBDBLINE=20196006; PubMed=10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D., RA Amatanteidae P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champs M., Pfeiffer B.D., RA Miklos G.L.G., Miklos G.L.G., RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Nelson C.R., Miklos G.L.G., Miklos G.L.G., RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D., RA Balow R.M., Bauu A., Baxendale J., Bayraktaroglu L., Beasley E.M., RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., RA Borovka D., Botchan M.R., Bouck J., Brokstein P., Brottier P., RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., RA Dobson K., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., RA Durbin R.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W., RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasper K., RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., RA Jalali M., Kalush F., Karpen G.H., Murphy L., Muzny D.M., Nelson D.L., RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacieb J.M., RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B., RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J., RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., RA Yeh J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; RA Science 287:2185-2195(2000). RA 12

RP FlyBase; FBgn0031735; CG11029.

DR InterPro; IPR001087; Lipase_GDSL.

DR Pfam; PF00657; Lipase_GDSL; 1.

DR PROSITE; PS01098; LIPOSE_GDSL_SER; 1.

SQ SEQUENCE 447 AA; 51281 MW; BBB3F53DCCB35DDB CRC64;

Query Match 6.6%; Score 515; DB 5; Length 447;

Best Local Similarity 31.1%; Pred. No. 1-3e-29; Mismatches 155; Indels 76; Gaps 14; Matches 137; Conservative 73; Mismatches 155; Indels 76; Gaps 14;

QY 293 SEDRQLDS-----TTLAWHJMRRMMPAGEKEPLSVKHGRPM 331

DB 46 SYDPRPENGIQXTDIDPDLRHLFLNTRQTLKALN-----TEAASSRGREG 95

QY 332 KCPSQESPVLFSYRSNSYLTRLQKQDKLVEREGABIRCPDKP-SDTVPTSYRLKPAD 390

DB 97 K-----LQAP---VSKKVPLCPTNVRSPSPPTSEHLRFG 131

QY 391 INVIGALGDSLITAGNGAGSSTPGNLTDVUOYRGHSWSYGGDENIGTVTLANTIREFPLS 450

DB 132 IDITAAGFDGLSISACNGIUS--ANNAIDMINEFRGLTFSGCGGLAMWRFVTLNPINKP 189

QY 451 LKGTSVGTGKETS-PNAFLQNAVAGGRAEGLPYOARRUVDLMDTRHFOEQMKILPF 509

DB 190 LYGAWSNLVNTINRSSLRPAIMRMSDRPLPQARVLDLDRHDMKRMWKLTVY 249

QY 510 IGGNDLICDFCNLDLVHY-SQNFTD---NIGKALDILHAEVPRAFVNLYTVLEVNLUDEL 564

DB 250 VGNNDI--CSDLGHWDTPQSFLQHARDLQRFLRHDVPRLLINIVPNIPVLST 306

QY 565 YQEKKVYCPMRMILSICPCEWLKFEDNSTLATALEEFNKFOEKTHOLIESGRYDTRFPT 624

DB 307 MTKVPLQCL-FVVRHRVGCHCLINDLRNRTFERNMDLTTRWQOLDMEARLPEF-RRQDFA 364

QY 625 VVQPFENNDIMPRTSEGUDPNSFAPDCFFHSSKSHRSRAASALMNMLEPVGOK---T 680

DB 365 IVAPMLTKLTAAPIPLDSDTWRFHOCFHFSORGHALISNLWNMSMLPDDQKPPSV 424

QY 681 TRHKPENKINTCQNQVQPL 701

DB 425 VPELFE---RVVCTAEPQL 442

RESULT 13
Q5Q01 PRELIMINARY; PRT; 460 AA.

AC Q5Q01; PRELIMINARY; PRT; 460 AA.

DT 01-DEC-2001 (TREMBrel. 19, Created)

DT 01-DEC-2001 (TREMBrel. 19, Last sequence update)

DE Hypothetical 51.9 kDa protein.

GN F13H8.11.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabdiidae; Peledorinae; Caenorhabditis.

OX NCBI_TAXID=239;

RN [1] SEQUENCE FROM N.A.

RC STRAIN-BRISCOL N2; MEDLINE=99069613; PubMed=9851916;

RA None; RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.", Science 28:2012-2018(1998).

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2; DING H.; RT "The sequence of C. elegans cosmid F13H8.1"; RL Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.

RA "The sequence of C. elegans cosmid F13H8.1"; DR "Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases." RN [3] SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2; RA Waterston R.; RT "Direct Submission." DR "Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases." RN EMBL; U23139; AAU02483.1; -.

DR WormPep; F13H8.11; CE2264.

DR InterPro; IPR001087; Lipase_GDSL.

DR Pfam; PF00657; Lipase_GDSL.

RESULT 14
Q21799 PRELIMINARY; PRT; 348 AA.

AC Q21799; PRELIMINARY; PRT; 348 AA.

DT 01-NOV-1996 (TREMBrel. 01, Created)

DT 01-NOV-1996 (TREMBrel. 01, Last sequence update)

DE R07B7.8 protein.

GN R07B7.8.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabdiidae; Peledorinae; Caenorhabditis.

OX NCBI_TAXID=6239;

RN [1] SEQUENCE FROM N.A.

RA Harris B.R.; MEDLINE=99069613; PubMed=9851916;

RA Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.

RA Sequence FROM N.A.

RC STRAIN-BRISTOL N2; DING H.; RT "Genome sequence of the nematode C.elegans: A platform for investigating biology.", Science 282:2012-2018(1998).

RA EMBL; 275955; CAB00118.1; -.

DR WormPep; R07B7.8; CE06270.

DR InterPro; IPR001525; C5_DNA_meth.

DR Pfam; PF00657; Lipase_GDSL.

DR PROSITE; PS0095; C5_MTASE_2.1.

DR Sequence 348 AA; 39489 MW; 91F5DC5139363E0 CRC64;

QY Query Match 5.2%; Score 406.5; DB 5; Length 348; Best Local Similarity 32.2%; Pred. No. 1e-21; Matches 117; Conservative 66; Mismatches 133; Indels 47; Gaps 11;

QY 1060 IKPAIENGSDFLCTEWKASNSV----PTSVHQLRPAIDIKVVALGDSLITAVGARN 113

RESULT 15

Db 15 VKHAKLE-----KVLNSIDPHEYPDYNMKEPHHRVIGAMGDSLTL--IGSRAE 61
 Qy 1114 NSSDLPTSWRGLSWSICGDGNLJETHTPNLLK---KENPVULFESTTWG-TAGLVN 1168
 Db 62 NTVGQORQYGPNAFTGMPDFEVDRHLYTNTIATAEKGNGKUFGGTGID-GENGTGLAV 121
 Qy 1169 AAEGRARADMPAQDWLVERMKNSPDINLEKDKWKLVTFIGND--LCHYCENPEAHLA 1225
 Db 122 AGGMKMSDILRQAKELYSRKANEKINENNDKVLVSLWIGTNVDTGIRLEDPIP-V 179
 Qy 1226 TETVOHQOALDITSEELPRAFYNUVNEMLASLYOSOGGKARMLAQONNCCLRSQSS 1285
 Db 180 DEYKSHIEKGKLYLKENLPRTIVSIVGMPAQLIQEAQ-----SILKNGKRA 226
 Qy 1286 LEQOELKUKNWMIQHGSFSFWHQQYTOR--EDPAVWQOPF-QNQNLTPNLERGDDLT 1341
 Db 227 RVEQNOKKLKD-DLSDGYRNVSDFDNNEHFNSNDFTWVOPATEYXTDYSREHKGNT 285
 Qy 1342 FNSEDCEHPSDRGHAEMAIALNMMLERPGRKITSNNFTHSRAKLUKCPSPESPYLYTLRN 1401
 Db 285 FVASDLEFLSKCCHAVLAKHYWNLFEPVGEGKTKRADLGDTKPKIVBLNEKLIKIVGN 345
 Qy 1402 SRL 1404
 Db 346 SKM 348

Search completed: January 6, 2004, 18:59:17
 Job time : 108 secs

P90862 PRELIMINARY; PRT; 528 AA.

ID P90862;
 AC P90862;
 DT 01-MAY-1997 (TREMBrel. 03, Created)
 DT 01-MAR-2003 (TREMBrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBrel. 23, Last annotation update)
 DE Hypothetical protein P36A2.9a.
 GN P36A2.9A.
 OS Caenorhabditis elegans.
 OC Cukarova; Metaxoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis;
 OX NCBI_TAXID:6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bernard N.;
 RL Submitted (Oct-1996) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9969613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C.elegans*: A platform for
 Science 283:2012-2018(1998).
 DR EMBL: 281077; CAB03072.2; -.
 DR Hypothetical protein.
 SQ SEQUENCE 528 AA; 59421 MW; 1B34202465B16261 CRC64;

Query Match 4.6%; Score 360.5; DB 5; Length 528;
 Best Local Similarity 29.9%; Pred. No. 5 5e-18; Mismatches 13;
 Matches 103; Conservative 59; Mismatches 134; Indels 49; Gaps 13;

Db 370 CPDKDPSTDTPVPSVHRUKPADINVIGALGDSLTAGNGAGSTGFTGNVLDVLTQYRGLSMVG 429
 Db 130 CRKTEBLTGTVGNPSPEDITIAAMGDAATGIGLWPNT---ADT--EHRGAISPIG 183
 Qy 430 GDENIGTVTLANILREBNPSIKGFSVGTGKETSPNAFLNQAVAGGRAEDLVRQARLVD 489
 Db 184 GSGTGGTGTINLIRESPKLVGSIQMGAL-PNHQINAVTGTATEDDUDQARRTR 242
 Qy 490 LMKNDTRHFOQDWKITHLFIGC3GNDLCDFCNLQDVHYSQPQFTNIGKLDIHAEVPRAF 549
 Db 243 RUKKLSLEDYHNEWIMMITIGTEELCSRCEGPSY-----DNIKAIKHLQIEHPKA- 294
 Qy 550 VNVIVTIVLIVNRELQEKKVCPRMILRSICPCVLFKEDDNTELATLIEFNKKFQF-- 606

THIS PAGE BLANK (uprto)